


***** (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Dec 9 08:36:34 1997; MasPar time 215.75 Seconds
Tabular output not generated. 980.525 Million cell updates/sec

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCCTTCATGGCCT.....AAAAAAGCGCGCCG 2027
Comp: CTTAAGCGGAAGTACCGGA.....TTTTTTTTTCGCGCGCG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0
Searched: 142080 seqs, 52183452 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq28
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 9.918; Variance 6.937; scale 1.430
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Description	ID	Pred. No.
1	201	9.9	Human gene signature	T23535	1.17e-97
2	88	4.3	Human Natriuretic Pep	Q10572	3.43e-32
3	84	4.1	Human Natriuretic Pep	Q10572	5.69e-30
4	44	2.2	Oligonucleotide probe	Q51746	8.73e-09
5	43	2.1	Base substituted E.co	N81164	2.71e-08
6	41	2.0	Oligonucleotide probe	Q51746	2.56e-07
7	40	2.0	Generic DNA sequence	Q70468	7.77e-07
8	40	2.0	Generic DNA sequence	Q70466	7.77e-07
9	40	2.0	Base substituted E.co	N81164	7.77e-07
10	39	1.9	Generic DNA sequence	Q70469	2.34e-06
11	38	1.9	Generic DNA sequence	Q70465	7.00e-06
12	38	1.9	Generic DNA sequence	Q70467	7.00e-06
13	39	1.9	Generic DNA sequence	Q70469	2.34e-06
14	38	1.9	Generic DNA sequence	Q70466	7.00e-06
15	38	1.9	Generic DNA sequence	Q70468	7.00e-06

16	36	1.8	114 12	070470	Generic DNA sequence	6.09e-05
17	35	1.8	114 12	070472	Generic DNA sequence	6.09e-05
18	36	1.8	114 12	070467	Generic DNA sequence	6.09e-05
19	34	1.7	114 12	070473	Generic DNA sequence	5.10e-04
20	34	1.7	114 12	070471	Generic DNA sequence	5.10e-04
21	34	1.7	114 12	070465	Generic DNA sequence	5.10e-04
22	34	1.7	114 12	070471	Generic DNA sequence	5.10e-04
23	34	1.7	114 12	070470	Generic DNA sequence	5.10e-04
24	32	1.6	114 12	070472	Generic DNA sequence	5.10e-04
25	30	1.5	565 6	Q35072	HCV envelope region n	4.09e-03
26	30	1.5	3871 2	N71302	HSV-1 gB and surround	3.13e-02
27	28	1.4	114 12	070473	Generic DNA sequence	3.13e-02
28	28	1.4	1454 1	N82107	PAP-I CDNA from human	2.27e-01
29	28	1.4	1466 1	N91353	Vascular anti-coagula	2.27e-01
30	28	1.4	1466 1	N80801	Sequence encoding vas	2.27e-01
31	28	1.4	1567 1	N81113	Plasmid contg. placen	2.27e-01
32	28	1.4	1575 1	N90112	Anticoagulant PP4 DNA	2.27e-01
33	28	1.4	1605 1	N91821	Endonexin II complete	2.27e-01
34	29	1.4	1840 11	Q65611	Feline zona pellucida	8.48e-02
35	26	1.3	31 17	Q99581	Human TPO anti-sense	1.54e+00
36	26	1.3	66 21	T13585	TSAR-9 library genera	1.54e+00
37	27	1.3	67 24	T14322	Primer used in the la	5.96e-01
38	27	1.3	82 21	T13610	DC43 TSAR library gen	5.96e-01
39	26	1.3	290 20	T21481	Human gene signature	1.54e+00
40	27	1.3	789 2	Q10792	Bovine vascular endot	5.96e-01
41	27	1.3	789 7	Q44255	Partial bovine VEGF-1	5.96e-01
42	27	1.3	829 2	Q10796	Bovine vascular endot	5.96e-01
43	27	1.3	961 2	Q10791	Bovine vascular endot	5.96e-01
44	27	1.3	961 7	Q44259	Bovine VEGF-164 codin	5.96e-01
45	26	1.3	1460 2	Q12679	PAP-I.	1.54e+00

ALIGNMENTS

RESULT 1
ID T23535 standard; CDNA to mRNA; 204 BP.
AC T23535;
DT 02-SEP-1996 (first entry)
DE Human gene signature HUMGS05382.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-AL.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 1393; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 204 BP; 54 A; 59 C; 35 G; 55 T;

Query Match 9.9%; Score 201; DB 20; Length 204;
Best Local Similarity 99.5%; Pred. No. 1.17e-97;
Matches 201; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 1 gatctcagctccggtccccaagcacactcctagctgctccagctcagctggcagct 60
QY 1207 GATCTCAGCTCCGTTCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCTGGCAGCT 1266
Db 61 tccccctgcttttgacgttttcaccccccagcattctcctgagttataagccacagag 120
QY 1267 TCCCCCTGCTTTTGACGCTTCATCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCTGGCAGCT 1326
Db 121 tggatagctgttttcacctaagaaagcccccagcattctgtagaataattcaact 180
QY 1327 TGGATAGCTGTTTTCACCTAAAGAAAGCCCAAGTCTTGTAGAAATATCAAACT 1386
Db 181 antaaatcatgaatatatttta 202
QY 1387 AATAAAATCATGAATATTTTA 1408

RESULT 2
ID Q10572 standard; DNA; 1047 BP.
AC Q10572;
DT 09-APR-1991 (first entry)
DE Human Natriuretic Peptide Receptor B.
KW NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW hyperaldosteronism; glaucoma; guanyl cyclase.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= signal sequence
FT Protein 12
FT /label= mature NPRB
FT Domain 23..455
FT /label= extracellular domain
FT /note= "binds natriuretic peptides A, B and C"
FT Domain 456..456
FT /label= transmembrane domain
FT Domain 479..1047
FT /label= cytoplasmic domain
FT /note= "GC and protien kinase activity"
FT /modified-site 24..26
FT /label= N-glycosite
FT /modified-site 35..37
FT /label= N-glycosite
FT /modified-site 161..163
FT /label= N-glycosite
FT /modified-site 195..197
FT /label= N-glycosite
FT /modified-site 244..246
FT /label= N-glycosite
FT /modified-site 277..279
FT /label= N-glycosite
FT /modified-site 349..351
FT /label= N-glycosite
FT /modified-site 600..602
FT /label= N-glycosite
FT WO9100292-A.
PD 10-JAN-1991.
PF 22-JUN-1990; U03586.
PR 23-JUN-1989; US-370673.
PA (GETH) GENENTECH INC.
PI Chang M, Goeddel D, Lowe D;
DR WPI; 91-036711/05.
DR N-PSDB; Q10324.
PT Natriuretic protein receptor B - for diagnosis and treatment of
PT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS Claim 3; Fig 1; 49pp; English.
CC The sequence was derived from the DNA encoding natriuretic peptide
CC receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC kinase activity. The DNA can be inserted into expression vectors
CC for the prodn. of the protein, opt. after being mutated to produce

CC NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr= 114,952). The protein (or variants) can be used in treatment of
CC natriuretic peptide disorders, and also to isolate peptides using
CC affinity chromatography. Antibodies with affinity for NPRB can
CC also be prepd.
SQ Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;
Query Match 4.3%; Score 88; DB 2; Length 1047;
Best Local Similarity 8.4%; Pred. No. 3.43e-32;
Matches 83; Conservative 278; Mismatches 618; Indels 8; Gaps 8;
Db 58 ranvndrvnssnnngacsnynannsvdnknyhdnndnnngngcvynaasvarnaswhrn 117
QY 667 GACATGCTTGAGTCGACCGCTTTCCCGAGACACGACCTTTGCATCCCCCTCGCTAGC 726
Db 118 nmntagasvsgnsakndhytrvrtgnsankngnnvntvnhghnnwtaraannndartdd 177
QY 727 AGCGAC-CACCTCCT-GCCAGCCACCGAGGAGCTCCAAAGGTATGTCAACCTCCGCAAA 784
Db 178 rnhyntangvnnannngsnsvnhvnyarnngnnnathnrrangrnyvncngnnmnnhnn 237
QY 785 AATAAAATGATGACACAGCAC-ATAATGAAACGGCTTTGTAAAAATGATTTT-GCACT 842
Db 238 nnanrnnntngdyvnyndvngnsnragntzratgrnvndarttrnnanrnnantvnt 297
QY 843 GAAATAAAAGTGAAGGAGATTAACCTACATCAACCGAGATACCAAAATCACTCTGGAGAC 902
Db 298 yrnnnnnnynnnnrnnrardngvngnsmnnnagcnydngnnnyannvntnnnggt 357
QY 903 CAAGAGCAAGACCATTTTACAAGCTGAACGGTGTGTCGAAAGGACCTGAAGAATCGGT 962
Db 358 rndgnrnnvnmgrryhgvtnvnmkndrn-tdnvnmwngdndsdnnnaahysgank 416
QY 963 GCTGTGGCTCAAAGACAGCTTGAGTGCACCTGTGAGGAGATGAACGACATCAACGCGCC 1022
Db 417 nnwvtgrnnnwkgannsdnnncandndnsdcktnntstnanvangtgnntnmngvsnn 476
QY 1023 CTATCTGCTCATGGGACAGAAACAGGGTGGGAGCTGCTCATCACCTCGGTGAAGCGGTG 1082
Db 477 nnrknmnnknnsnnwrrnnnnngnsnryhkgagrntnsnrgssyngsmthgkynn 536
QY 1083 GCAGAGGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGATCCGCAAGTGCAGTGTCTA 1142
Db 537 nantghnkgvnnvankhvkkrnntrnnvnnnnkhmrdrnnnnhtrnnngacndnnnnvnt 596
QY 1143 GTCCCGGCATCCTGATGGCTCCGACAGCGCTGTCTCCAGAGCAGCGCTGACCATTTCTGT 1202
Db 597 nycnrgsnndnnndsmnrysnndndvngmnnhnnshsngsknsncvvdsm 656
QY 1203 -CCGGGATCTCAGCTCCGCTCCCAAGCACACTCTCTAGCTGCTCCAGCTCAGCCTGGG 1261
Db 657 vnkntdygnasnrstannddnnanyakkntannnnsgnnnnnttgmaadvyngnnnnn 716
QY 1262 CAGCTTCCCCCTGCTTTTGACAGTTTGACATCCCGACATTTCTCTGAGTTATAAGGCCAC 1321
Db 717 nanrs-gnnyngndnsnknvnnvkrngnrynrynsndrtnnnnnvnnnmrncwandnan 775
QY 1322 AGGAGTGGATAGCTGTTTTCACCTAAAGGAAAGCCCAAGTCTGTAGAATATTC 1381
Db 776 rndngnknknrrnnknnggtsnndnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 834
QY 1382 AAACTAATAAAATCATGAATATTTTATGAAGTTTAAATAATAGCTCACTTAAAGCTAGT 1441
Db 835 ynnnnhsvannnkrgntvnanandsvtynsdvngtansanstnmvntnnndnycnd 894
QY 1442 TTTGAATAGTGTCAACTGTGACTTGGGTCTGGTGTGTTGTTGTTTGTAGTCAG 1501
Db 895 anndndvkvntngdaymvvsgnngnrgnrrhannnnmarmananddvssnrnrhndnn 954
QY 1502 CTGATTTTCACTTCCCACTGAGGTGTCTAATAACATGCAAAATGCTTCAATTTCTCTGTG 1561
Db 955 rnrngvhtgncvagvgnknmrycnnngdvtntasrmnsgnngnknhvssttkdandnnc 1014


```

Db 119 rsvkydcynachddhyvbybbvynvhnncnccbnhvnchvhnhrnwayvrhd 178
Cp 283 GCAGCAGCAGCAGGCGCTCGAGCATCTGCGCGCGGACCCCGAGGGGCGAGAGGA 224
Db 179 avrddvhc 186
Cp 223 GCGGAGCC 216

RESULT 10
ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'z'; z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994. U00977.
PF 01-FEB-1994; US-013416.
PR 01-FEB-1993; US-176500.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UUNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Key BK;
DR WPI; 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 253pp; English.
CC Q70469 is a generic DNA sequence used to generate random TSAR peptide
CC This generic formula can be represented as follows: X(TGC)(NNB)10-
CC (TGC)(NNB)6Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC sites (X is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in Q70465-68. Other specific peptides
CC generated by these generic sequences are shown in R65150-54. TSARs are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active. They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARs or compsns. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 G; 4 C; 4 T;

Query Match 1.9%; Score 39; DB 12; Length 114;
Best Local Similarity 7.2%; Pred. No. 2.34e+06;
Matches 8; Conservative 32; Mismatches 71; Indels 0; Gaps 0;

Db 1 tgcnnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn 60
Qy 595 TGCCACTGCTGCTGCGTGGAGTGGAGCCGCTGCGCCCGGTCATGCTCGGCTTCGGC 654
Db 61 nnbnbnbtgcnnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn 111
Qx 655 TTCCCTGCGCGGACATGCTTGAGTGCAGCGGTTTCCCGCAGGACACGAC 705

```

```

RESULT 11
ID Q70465 standard; DNA; 114 BP.
AC Q70465;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'z'; z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994. U00977.
PF 01-FEB-1994; US-013416.
PR 01-FEB-1993; US-176500.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UUNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Key BK;
DR WPI; 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 253pp; English.
CC Q70465 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)6(TGC)(NNB)11Z(TGC)(NNB)3Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed
CC activity allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.9%; Score 38; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 7.00e+06;
Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;

Db 3 bnnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn 62
Qy 273 CTGCTGCTGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 332
Db 63 bnnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn 114
Qy 333 CCTCTTTGGCCAGCCGACTTCTCTACAGCGCAGCAATTGCAAGCCCATC 384

RESULT 12
ID Q70467 standard; DNA; 114 BP.
AC Q70467;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.

```



```
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK.
DR WPI; 94-279739/34.
DR P-PSDB; R65152.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70466 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)10(TGC)2(NNB)4Z(NNB)8(TGC)(NNB)
CC -91. X and Y are flanking restriction sites (X is not the same as Y)
CC that are not specified further. Other generic sequences are shown in
CC Q70466-68. Other specific peptides generated by these generic sequences
CC are shown in R65151-54. TSARs are concatenated heterofunctional proteins
CC or peptides, comprising at least two functional regions - a binding
CC domain with affinity for a ligand and a second effector peptide portion
CC that is chemically or biologically active. They may further comprise a
CC linker peptide between the 2 domains. The oligonucleotides are also
CC designed so that the expressed peptide contains 2 or 4 cysteine residues
CC positioned in, or flanking, the unpredicted or variant residues. These
CC residues confer some degree of conformational rigidity to the peptides.
CC The TSARs or compns. comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety, eg. metal
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC on the cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 1.9%; Score 38; DB 12; Length 114;
Best Local Similarity 8.9%; Pred. No. 7.00e-06;
Matches 10; Conservative 31; Mismatches 71; Indels 0; Gaps 0;

Db 3 btgcnbnbnbnbnbnbnbnbnbnbnbnbtgctgcnbnbnbnbnbnbnbnbnbnbnbn 62
Cp 513 CTGCTTCATGACCGGGGATCCAGCGCGCGCTGCTCCAGCACCTTCATGCTC 454
Db 63 bnnbnbnbnbnbnbnbnbnbnbnbnbtgcnbnbnbnbnbnbnbnbnbnbnbnbn 114
Cp 453 GTGGCCGACGAGTGGCGAGCGCATGCTCTGTATTGATGCGGTGGCAC 402

RESULT 15
ID Q70468 standard; DNA; 114 BP.
AC Q70468;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYN-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK.
DR WPI; 94-279739/34.
DR P-PSDB; R65154.
PT Identifying proteins or peptide(s) which bind a ligand - by
```

```
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)6Z(NNB)7(TGC)(NNB)10Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.9%; Score 38; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 7.00e-06;
Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;

Db 3 bnnbnbnbnbnbnbnbnbnbnbnbnbtgcnbnbnbnbnbnbnbnbnbnbnbnbnbn 62
Cp 543 GCACAGGAACCTCTTGGTGTCGGGTGGCACTGCTTCATGACCGGGATCCAGCGCC 484
Db 63 bnnbnbnbnbnbnbnbnbnbnbnbnbtgcnbnbnbnbnbnbnbnbnbnbnbnbnbn 114
Cp 483 GGCCTGCTCCAGCACCTCTTCATGCTGCTGCGCCGACGAGTGGGCAGC 432

Search completed: Tue Dec 9 08:43:11 1997
Job time : 397 secs.
```

MORF

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPPerch_nnn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Dec 9 09:11:05 1997; MasPar time 470.39 Seconds
Tabular output not generated. 965.771 Million cell updates/sec

Title: >848439-2-trans
Description: (1-885) from translate.seq
Perfect Score: 734
N.A. Sequence: 1 ATGTYNCARGNCCNGGNSW.....SNATNMGAAATYTCARTGY 885
Comp: TACRANGTYCCNGGNCNSW.....WNTANKNTYRANGTYACR

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 707517 seqs, 256659390 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-SWS

1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99

EST-SWS-TWO

100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
170:EST170 171:EST171 172:EST172 173:EST173 174:EST174

Database:

175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
190:EST190 191:EST191 192:EST192 193:EST193 194:EST194
195:EST195 196:EST196 197:EST197 198:EST198

Statistics: Mean 15.226; Variance 4.441; scale 3.429

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	345	47.0	462	174	W77968	zd70ell.r1	Soares fet	0.00e+00
2	319	43.5	396	178	AA073862	mm97f07.r1	Stratagene	0.00e+00
3	304	41.4	426	72	H87071	ys74d07.r1	Homo sapie	0.00e+00
4	279	38.0	406	61	H16121	yl28c05.r1	Homo sapie	0.00e+00
5	277	37.7	414	156	AA024771	ze76g10.r1	Soares fet	0.00e+00
6	242	33.0	344	61	H14917	yl2ed03.r1	Homo sapie	5.47e-272
7	210	28.6	299	148	W08345	mb41f02.r1	Soares mou	1.74e-227
8	163	22.2	231	65	H29095	ym31e03.r1	Homo sapie	4.01e-163
9	151	20.6	250	172	AA061047	mj78a12.r1	Soares mou	5.97e-147
10	63	8.6	531	12	M89402	CEL08B2	Caenorhabdit	2.81e-35
11	57	7.8	360	141	C09584	C.elegans	CDNA clone	1.49e-28
12	52	7.1	360	89	CELK117B4F	C.elegans	CDNA clone	3.90e-23
13	41	5.6	360	89	CELK079D1F	C.elegans	CDNA clone	5.43e-12
14	40	5.4	256	47	H45312	yo65b03.r1	Homo sapie	4.79e-11
15	37	5.0	206	101	N63268	yz88d01.s1	Homo sapie	2.72e-08
16	37	5.0	418	101	N62714	yz76d05.s1	Homo sapie	2.72e-08
17	36	4.9	341	77	CELK119F4R	C.elegans	CDNA clone	2.10e-07
18	34	4.6	345	129	HSC0NC092	H. sapiens	partial CD	1.11e-05
19	34	4.6	336	123	HSBA0B082	H. sapiens	partial CD	1.11e-05
20	33	4.5	227	184	AA017479	ze38h10.r1	Soares ret	7.63e-05
21	33	4.5	350	187	AA144540	mr17f12.r1	Soares mou	7.63e-05
22	33	4.5	414	194	AA169546	z089b08.r1	Stratagene	7.63e-05
23	33	4.5	428	95	N41923	yy07a02.r1	Homo sapie	7.63e-05
24	33	4.5	433	194	AA169521	z089a08.r1	Stratagene	7.63e-05
25	33	4.5	443	113	W03474	z08a10.r1	Soares mel	7.63e-05
26	33	4.5	458	23	R42865	yg06a10.s1	Homo sapie	7.63e-05
27	32	4.4	219	129	HSC0EA112	H. sapiens	partial CD	4.99e-04
28	32	4.4	226	156	AA027879	zk05e04.r1	Soares pre	4.99e-04
29	32	4.4	242	106	HSC39E102	H. sapiens	partial CD	4.99e-04
30	32	4.4	242	119	W44804	z098g09.s1	Soares par	4.99e-04
31	32	4.4	243	133	N84441	KR8808F	Homo sapiens	4.99e-04
32	32	4.4	253	135	R58334	G2318	Fetal heart Hom	4.99e-04
33	32	4.4	262	81	T35329	EST82912	Homo sapiens	4.99e-04
34	32	4.4	345	105	HSC1EE122	H. sapiens	partial CD	4.99e-04
35	32	4.4	384	95	N42176	yy07ell.r1	Homo sapie	4.99e-04
36	32	4.4	453	95	N42436	yy09h05.r1	Homo sapie	4.99e-04
37	32	4.4	473	127	W90442	zh78c02.s1	Soares fet	4.99e-04
38	32	4.4	584	4	T67709	yc27c11.s1	Homo sapie	4.99e-04
39	31	4.2	252	109	HUM069H03B	Human fetal brain	CDN	3.11e-03
40	31	4.2	293	22	R38715	yd03b08.s1	Homo sapie	3.11e-03
41	31	4.2	295	26	R32481	yg80c12.s1	Homo sapie	3.11e-03
42	31	4.2	362	24	R46624	yg53c02.s1	Homo sapie	3.11e-03
43	31	4.2	372	72	H87471	vw17f06.s1	Homo sapie	3.11e-03
44	31	4.2	390	1	CELK025G2R	C.elegans	CDNA clone	3.11e-03
45	31	4.2	483	98	N52277	yv47el2.s1	Homo sapie	3.11e-03

ALIGNMENTS

1 W77968 462 bp mRNA EST 17-OCT-1996
LOCUS zd70ell.r1 Soares fetal heart NBH19W Homo sapiens CDNA clone
DEFINITION 346028 5'
ACCESSION W77968
NID g1388502
KEYWORDS EST.

/sex-hermaphrodite

адрес: г. Москва, ул. Мухоморова, д. 10, стр. 1

ORIGIN

W O R L D

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 09:33:55 1997; MasPar time 244.79 Seconds
Tabular output not generated. 993.543 Million cell updates/sec

Title: >848439-2-trans
Description: (1-885) from translate.seq
Perfect Score: 734
N.A. Sequence: 1 ATGYTNCARGGCCNGGNS.....SNATNMGNAARTNCARTGY 885
Comp: TACRANGTYCCNGGCCNSW.....WNTAKNCNTYRANGTYACR

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 359085 seqs, 137405154 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS-THREE
1: EST199 2: EST200 3: EST201 4: EST202 5: EST203 6: EST204
7: EST205 8: EST206 9: EST207 10: EST208 11: EST209 12: EST210
13: EST211 14: EST212 15: EST213 16: EST214 17: EST215
18: EST216 19: EST217 20: EST218 21: EST219 22: EST220
23: EST221 24: EST222 25: EST223 26: EST224 27: EST225
28: EST226 29: EST227 30: EST228 31: EST229 32: EST230
33: EST231 34: EST232 35: EST233 36: STS1 37: STS2 38: STS3
39: STS4 40: STS5 41: STS6 42: STS7 43: STS8 44: STS9 45: STS10
46: STS11 47: STS12 48: STS13
EST-STS-FOUR
49: gNEST1 50: gNEST2 51: gNEST3 52: gNEST4 53: gNEST5
54: gNEST6 55: gNEST7 56: gNEST8 57: gNEST9 58: gNEST10
59: gNEST11 60: gNEST12 61: gNEST13 62: gNEST14 63: gNEST15
64: gNEST16 65: gNEST17 66: gNEST18 67: gNEST19 68: gNEST20
69: gNEST21 70: gNEST22 71: gNEST23 72: gNEST24 73: gNEST25
74: gNSTs 75: enEST1 76: enEST2 77: enEST3 78: enEST4
79: enEST5 80: enEST6 81: enEST7 82: enEST8 83: enEST9
84: enEST10 85: enEST11 86: enEST12 87: enEST13 88: enEST14
89: enEST15 90: enEST16 91: enEST17 92: enEST18 93: enEST19
94: enEST20 95: enEST21 96: enEST22 97: enEST23 98: enEST24
99: enEST25 100: enEST26 101: enEST27 102: enEST28
103: enEST29 104: enEST30 105: enEST31 106: enSTs 107: ueEST1
108: ueEST2

Database:

Statistics: Mean 15.318; Variance 4.699; scale 3.259

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result % Query

No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	357	48.6	458	33	AA105749	m184a03.r1 Stratagene	0.00e+00
2	319	43.5	396	54	AA073862	mm57f07.r1 Stratagene	0.00e+00
3	235	32.0	394	101	MMAA60087	va52q08.r1 Soares mou	4.81e-249
4	210	28.6	299	94	NW3451	mb41f02.r1 Soares mou	4.84e-216
5	148	20.2	184	20	AA209833	mo90h01.r1 Beddington	1.07e-135
6	145	19.8	523	56	AA122822	mr03c11.r1 Soares mou	6.95e-132
7	145	19.8	523	6	AA122822	mr03c11.r1 Soares mou	1.67e-120
8	136	18.5	287	32	AA082155	ze88f06.r1 Soares fet	5.43e-25
9	55	7.5	502	21	AA020088	mh49d08.r1 Soares mou	1.05e-08
10	38	5.2	213	28	AA023595	mh80d03.r1 Soares mou	1.05e-08
11	37	5.0	421	82	HS147080	zr18a11.sl Stratagene	7.52e-08
12	37	5.0	421	63	AA227473	zr18a11.sl Stratagene	7.52e-08
13	36	4.9	512	29	AA044908	zf51e10.r1 Soares ret	5.22e-07
14	35	4.8	432	29	AA030216	mh87g08.r1 Soares mou	3.49e-06
15	35	4.8	519	39	G08492	human STS CHLC.ATA11D	3.49e-06
16	34	4.6	288	36	DM57F11T	D. melanogaster STS d	2.25e-05
17	34	4.6	496	10	AA013810	mh06e01.r1 Soares mou	2.25e-05
18	33	4.5	241	21	AA019228	ze59c03.r1 Soares ret	1.39e-04
19	33	4.5	252	37	G03067	human STS WI-5108	1.39e-04
20	33	4.5	350	56	AA144540	mr17f12.r1 Soares mou	1.39e-04
21	32	4.4	345	44	G27227	human STS SHGC-32136	8.23e-04
22	32	4.4	552	12	AA180957	zp41c03.r1 Stratagene	8.23e-04
23	31	4.2	337	10	AA009696	ze82h05.sl Soares fet	4.66e-03
24	31	4.2	407	89	HSAA62672	zs23f08.r1 Soares NDH	4.66e-03
25	31	4.2	417	6	AA122184	zn28c08.sl Stratagene	4.66e-03
26	31	4.2	444	17	AA197092	zg10e04.r1 Stratagene	4.66e-03
27	31	4.2	451	39	G09558	human STS CHLC.GCT8C0	4.66e-03
28	31	4.2	491	25	N53936	yv59d10.sl Soares fet	4.66e-03
29	31	4.2	516	12	AA181149	zp59c10.sl Stratagene	4.66e-03
30	30	4.1	177	36	DM52H6T	D. melanogaster STS d	2.51e-02
31	30	4.1	213	10	AA018016	mh45c07.r1 Soares mou	2.51e-02
32	30	4.1	278	105	TB2692	T1343 MVAT4 bloodstre	2.51e-02
33	30	4.1	305	21	AA019656	ze61f11.sl Soares ret	2.51e-02
34	30	4.1	339	43	G26024	human STS EST186690	2.51e-02
35	30	4.1	402	1	N97824	1353C3 czappFDD2.1, D	2.51e-02
36	30	4.1	407	21	AA022630	ze71h06.sl Soares fet	2.51e-02
37	30	4.1	425	43	G24362	human STS WI-13950	2.51e-02
38	30	4.1	435	34	AA141629	CK02453.3prlme drosop	2.51e-02
39	30	4.1	440	26	N64746	yz30a03.sl Soares mul	2.51e-02
40	30	4.1	455	62	AA224647	CPEST.119 psk1minusc	2.51e-02
41	30	4.1	455	80	CPAA24647	CPEST.119 psk1minusc	2.51e-02
42	30	4.1	471	44	G27881	human STS SHGC-34074	2.51e-02
43	30	4.1	503	70	AA241149	mv26g08.r1 GuayWoodfo	2.51e-02
44	30	4.1	503	92	NW1157919	mv26g08.r1 GuayWoodfo	2.51e-02
45	30	4.1	527	9	AA131693	z134f05.sl Soares pre	2.51e-02

ALIGNMENTS

RESULT	1
LOCUS	AA105749 458 bp mRNA EST 04-FEB-1997
DEFINITION	m184a03.r1 Stratagene mouse kidney (#937315) Mus musculus CDNA clone 518668 5' similar to TR:G1151260 G1151260 TRANSMEMBRANE RECEPTOR. ;.
ACCESSION	AA105749
NID	G1654838
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 458) Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE	The WashU-HMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:312516

Seq primer: -28ml3 rev1 ET from Amersham

High quality sequence stop: 457.

Location/Qualifiers

FEATURES

source

1..458
/organism="Mus musculus"
/strain="C57/B16"
/note="vector: pBluescript SK-; Site.1: EcoRI; Site.2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
sequence: 5' GAATTCGGCAGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3'"
/clone="518668"
/clone_lib="Stratagene mouse kidney (#937315)"
/sex="females"
/dev_stage="4 weeks"
/lab_host="SOLR (kanamycin resistant)"
<1..>458

BASE COUNT 119 a 140 c 116 g 83 t

ORIGIN

Query Match 48.6%; Score 357; DB 33; Length 458;
Best Local Similarity 60.6%; Pred. No. 0.00e+00;
Matches 280; Conservative 103; Mismatches 75; Indels 4; Gaps 4;
Db 1 agcaggggcgatgattccgctggttcataagacagtcgacccgagacacagaagt 60
QY 221 ARCARGCGNGCNGTGGATNCCTTNGTNGATGAACARTGYCAYCCNGAYACAAARAAT 280
Db 61 tccgtgctcgtcttcgcccctgctgtctcgacagcctagatgagaccaccagccgt 120
QY 281 TTYTGTGWSNYTNTTYGCNCNGTNGTYTNGAYTNGAYTNGAYACNATNCARCCNT 340
Db 121 gtcaactcgtctg-gtgcaggtgaaggaccgctgctgcccgcgtcattcgcctcgtc 178
QY 341 GYCAWSNYTNTGYTNCARGTNAARGAYMGNTGYGCNCNGTNGATGWSNGCNTTYGGNT 400
Db 179 tccctggccagacatgctggagtcgacccgttcccgag-acaaacacctctcatcc 237
QY 401 TYCCNTGCCNGAYATGYTNGARTGYGAYMGNTTYCCNCARGAYAYGAYTNTGYATNC 460
Db 238 ccctcgtctagtag-gaccacctcctcgccgcccacagagagaagctcccaaggtgtgaag 296
QY 461 CNYTNGCWSNWSNGAYCAYTNTYCCNCNGCNCARGARGCNCNNAARGTNTGYGARG 520
Db 297 cctgcaaaaccaaagaatgaggacgaacacgacatcatcattgaaacccctttgtaaaatgact 356
QY 521 CNTGYAARAAYAAAYGAYGAYAYGAYATNATGARGACNTNTGYAARAAYGAYT 580
Db 357 tcgcaactgaaatacaagtgaggagatacgttacatcaacagagacacacagatcatcc 416
QY 581 TYGCNTYTAARATNAARGTNAARGARATNACNTATATNAAYMGNGAYACNNAARATNATNY 640
Db 417 tggagacaaagacagaccatttacaagctgaacggcggtg 458
QY 641 TNGARACNAARWSNAARACNATNTATAYATNTAAYTNGGNGTNW 682

RESULT 2 AA073862 396 bp mRNA EST 15-FEB-1997
LOCUS mn97f07.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
DEFINITION 536389 5' similar to TR:G1151260 TRANSMEMBRANE RECEPTOR.

ACCESSION AA073862

NID

g1595602

EST.

house mouse.

SOURCE

Mus musculus

ORGANISM

Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata;

Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Mus.

1 (bases 1 to 396)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

The WashU-HHMI Mouse EST Project

Unpublished (1996)

TITLE

JOURNAL

COMMENT

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:323325

Seq primer: -28ml3 rev1 ET from Amersham

High quality sequence stop: 390.

FEATURES

source

1..396
/organism="Mus musculus"
/strain="NIH/Swiss"
/note="vector: pBluescript SK-; Site.1: EcoRI; Site.2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. 93
pooled NIH/Swiss 13 day embryo hearts. Average insert
size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5'
GAATTCGGCAGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3'"
/clone="536389"
/clone_lib="Stratagene mouse heart (#937316)"
/sex="pooled"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
<1..>396

mRNA

BASE COUNT 93 a 126 c 103 g 74 t

ORIGIN

Query Match 43.5%; Score 319; DB 54; Length 396;

Best Local Similarity 61.7%; Pred. No. 0.00e+00;

Matches 245; Conservative 88; Mismatches 62; Indels 2; Gaps 2;

Db 1 agcaggcgggcg-cggattccgctggttcataagacagtcgacccgagacacagaagt 59

QY 221 ARCARGCGNGCNGTGGATNCCTTNGTNGATGAACARTGYCAYCCNGAYACNAAARAAT 280

Db 60 tccgtgctcgtcttcgcccctgctgtctcgacagcctagatgagaccatccagcgt 119

QY 281 TTYTGTGWSNTNTTYGCNCNGTNGTYTNGAYTNGAYTNGAYACNATNCARCCNT 340

Db 120 gtcaactcgtctg-gtgcaggtgaaggaccgctgctgcccgcgtcattcgcctcgtcgt 178

QY 341 GYCAWSNYTNTGYTNCARGTNAARGAYMGNTGYGCNCNGTNGATGWSNGCNTTYGGNT 400

Db 179 tccctggccagacatgctggagtcgacccgttcccgagagacacacctctcatcc 238

QY 401 TYCCNTGCCNGAYATGYTNGARTGYGAYMGNTTYCCNCARGAYAYGAYTNTGYATNC 460

Db 239 ccctcgtctagtagaacacacctcctcgccggccacagagaagctcccaaggtgtggaag 298

QY 461 CNYTNGCWSNWSNGAYCAYTNTYCCNCNGCNCARGARGCNCNNAARGTNTGYGARG 520

Db 299 cctgcaaaaccaaagaatgaggagcgaacacgacatcatggaacccctttgtaaaatgact 358

```

QY 521 CNTGYAARAARAAYGAYGAYAAAYGAYATNATGGARACNTYNTGYAARAAYGAYT 580
Db 359 tgcactgaaatcaaaagtgaagagataacgtacat 395
QY 581 TYGCNTNAARATNAARGTNAARGARATNACNTAYAT 617

RESULT 3
ID MAA60087 standard; RNA; EST; 394 BP.
AC AA260087;
NI g1896590
DT 19-MAR-1997 (Rel. 51, Created)
DE va52g08.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 735038 5'.
KW EST.
OS Mus musculus (house mouse)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Mus.
RN [1]
RP 1-394
RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA Waterston R.;
RA "The WashU-HHMI Mouse EST Project";
RL Unpublished.
CC Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC Washington University School of MedicineP 4444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: mouseest@wustl.edu This clone is available
CC royalty-free through LLNL ; contact the IMAGE Consortium
CC (info@image.llnl.gov) for further information. MGI:452086 Seq
CC primer: -28ml3 rev2 ET from Amersham.
FH Location/Qualifiers
FH source
FH 1..394
FH /organism="Mus musculus"
FH /strain="C57BL/6J"
FH /note="Vector: pT73D-pac (Pharmacia) with a modified
FH polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FH was primed with a Not I - oligo(dT) primer [5,
FH TGTACCAATCTGAAGTGGGAGCGCGCTATTATTTTITTTT 3'], o
FH total mouse RNA [provided by Minoru Ko, Wayne State Univ.];
FH double-stranded cDNA was ligated to Eco RI adaptors
FH (Pharmacia), digested with Not I and cloned into the Not I
FH and Eco RI sites of the modified pT73 vector. Library wen
FH through one round of normalization, and was constructed by
FH Bento Soares and M. Fatima Bonaldo."
FH /clone="735038"
FH /dev_stage="12.5dpc total fetus"
FH /sex="unknown"
FH /lab_host="DH10B"
FH <1..>394
FH mRNA
FH Sequence 394 BP; 104 A; 96 C; 118 G; 76 T; 0 other;

Query Match 32.0%; Score 235; DB 101; Length 394;
Best Local Similarity 57.5%; Pred. No. 4.81e-249;
Matches 164; Conservative 71; Mismatches 50; Indels 0; Gaps 0;

Db 1 aaggagataacgtacatcaacagacacacagatctctggagacaagagacacc 60
QY 601 AARGARATNACNTAYATNAAYMGNGAYACNAARATNATNTNGARACNAARACN 660
Db 61 atttacaagctgaacggcggtccgaaggacctgaagaatccggtggtcctaaa 120
QY 661 ATNTAARYTNAAYGGNGTWSNGARMGNGAYTNAARARWSNGTNTTGGYTNAAR 720
Db 121 gacagcctgcagtgcaacctgtgaggagatgaacgacatcaacgctcgtatcgtgctatg 180

```

```

QY 721 GAYWSNTNCACTGACNTGIGARGARATGAAYGAYATNATNAAYGCNCNTAYTNGTATG 780
Db 181 ggacagaagcagggcgagctggtgatcacctccctggaacggtggaagggccag 240
QY 781 GGCARAAARCAGGCGNGARYTNGTATNACNWSNCTNAARMGNTGGCARAARGNCAR 840
Db 241 agagagttcaagcgcatctccgcgacgcatccgcaagctgcaatgc 285
QY 841 MNGARTTYAARMGNATNWSNMGNWSNATNMGNNAARTNCAATGY 885

RESULT 4
ID MM3451 standard; RNA; EST; 299 BP.
AC W08345;
NI g1282366
DT 27-APR-1996 (Rel. 47, Created)
DE 05-MAR-1997 (Rel. 51, Last updated, Version 2)
DE mb41f02.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 331995
DE 5'.
KW EST.
OS Mus musculus (house mouse)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Mus.
RN [1]
RP 1-299
RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA Waterston R.;
RA "The WashU-HHMI Mouse EST Project";
RL Unpublished.
CC Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC Washington University School of MedicineP 4444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: mouseest@wustl.edu This clone is available
CC royalty-free through LLNL ; contact the IMAGE Consortium
CC (info@image.llnl.gov) for further information. MGI:213395 Seq
CC primer: mob.REGA+ET High quality sequence stop: 280.
FH Location/Qualifiers
FH source
FH 1..299
FH /organism="Mus musculus"
FH /note="Vector: pT73D (Pharmacia) with a modified
FH polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FH was primed with a Not I - oligo(dT) primer [5,
FH TGTACCAATCTGAAGTGGGAGCGCGCATTTTITTTTITTTT 3'],
FH double-stranded cDNA was size selected, ligated to Eco RI
FH adaptors (Pharmacia), digested with Not I and cloned into
FH the Not I and Eco RI sites of a modified pT73 vector
FH (Pharmacia). Library went through one round of
FH normalization to a Cot = 5. Library constructed by Bento
FH Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr
FH Minoru Ko (Wayne State University).
FH /clone="33195"
FH /dev_stage="19.5 dpc total fetus"
FH /lab_host="DH10B (ampicillin resistant)"
FH <1..>299
FH mRNA
FH Sequence 299 BP; 79 A; 95 C; 72 G; 53 T; 0 other;

Query Match 28.6%; Score 210; DB 94; Length 299;
Best Local Similarity 59.5%; Pred. No. 4.84e-216;
Matches 175; Conservative 68; Mismatches 46; Indels 5; Gaps 5;

Db 11 gctcgtcactcgtct-cgtgcag-tgaaggaccgtgcgc-cggtcatgtccgctt 67
QY 336 RCCNTGYCAWSNTNTGYGTNCARNAARGAYTMGNTGNCNCNTGNTGWSNGNTT 395
Db 68 cg-cttccccctggcagacatgctggagtcgacgcttctccgcgag-acaacgacctctg 125

```


[illegible][illegible]


```

/lab_host="DH10B (ampicillin resistant)"
/dev_stage="adult"
/lab_host="DH10B"
<1..>287
BASE COUNT      62 a      84 c      85 g      51 t      5 others
ORIGIN

Query Match          18.5%; Score 136; DB 32; Length 287;
Best Local Similarity 58.6%; Pred. No. 1.67e-120;
Matches 102; Conservative 41; Mismatches 30; Indels 1; Gaps 1;

Db    1   ggctcaagacagcttgcagtcgacctggaggatgaacacatcaacgccctc 60
QY    713  GGYNARGATWSNYTNCARTGYACNTGCGARATGAAGAYATNAAYGCNCTA 772
Db    61   tggctcagggcacaaacaggggtggggagctgggatccctcggtgaagcggtggcaga 120
QY    773  TNGTNATGGNCARAARCARGGNGGARVYTGTTGNATNACNWSGTNAARMGNTGGC 832
Db    121  aggggcagagagagttaaacgcattcncggngnngnatcccgaagctgcagtgc 174
QY    833  ARGNCARMGNGARTTYAARMGNATNWSMGNW-SNATNMGNARYTNCARTGY 885

RESULT  9     AA020088      502 bp      mRNA      EST      21-JAN-1997
LOCUS    mH49D08.r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
DEFINITION clone 45839 5' similar to PIR:A45054 A45054 probable intercellular signal transducer or transmitter Fz-1 ;.
ACCESSION AA020088
NID       G1483933
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 502)
AUTHORS  Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE     The WashU-HMI Mouse EST Project
JOURNAL   Unpublished (1996)
COMMENT   Contact: Marra M/Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:271175
Possible reversed clone: similarity on wrong strand
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 444.
Location/Qualifiers
1..502
/organism="Mus musculus"
/strain="C57BL/6J"
/note="vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAACTGAACTGGAGCGGCCGGGAATTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
/clones="445839"
/clone_lib="Soares mouse placenta 4NDMP13.5 14.5"
/sex="unknown"

FEATURES             source
source               1..502
location              1..502

```

```
T 3'; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M.Fatima Bonaldo."
/clone="457253"
/clone_lib="Soares mouse placenta 4NDMP13.5 14.5"
/sex="unknown"
/dev_stage="adult"
/lab_host="DH10B"
<1..>213
BASE COUNT      43 a  51 c  75 g  44 t
ORIGIN

Query Match      5.2%; Score 38; DB 28; Length 213;
Best Local Similarity 36.6%; Pred. No. 1.05e-08;
Matches 74; Conservative 30; Mismatches 98; Indels 0; Gaps 0;

Db 11 gatggcgacggtgtagcggtgtagtcaggcaacagatggcggtgtacatag 70
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 349 RNSRTGRCANGVYTGATNGTYTCRTCNARRTCRTCNARRCANACNGGCRANARNW 290
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 agcacgggaagcgcaggtcggtgagcagtggtatctccacaagcgccagaattggt 130
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 289 SRCANARRAAYTTTNGTTCGGRGTCAYTGTTCATNACNARNGGNATCCAGCNC 230
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 131 gcactccagcgctctcgtcggtcgtgcatggtggaactggtggcgatggtcgctca 190
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 229 CNGCYGTCNARNACYTCYTTCATNGTYTCRTGNCNARNARRTTNGGNARCKCATRT 170
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 191 ggttgtagcgatgcctcgcca 212
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Cp 169 TYTGRTAYTCNATNCCRTGRCA 148
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
ID HS1147060 standard; RNA; EST; 421 BP.
AC AA227473;
NI g1849027
DT 27-FEB-1997 (Rel. 51, Created)
DE zrl8all.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
DE cDNA clone 663740 3'.
KW EST.
OS Homo sapiens (human)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP 1-421
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevaskis E., Waterston R., Williamson A., Wohlmann P., Wilson R.;
RA "The WashU-Merck EST Project";
RL Unpublished.
CC Contact: Willson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC estewatson.wustl.edu This clone is available royalty-free through
CC LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for
CC further information. Seq primer: -41ml3 fwd. ET from Amersham High
CC quality sequence stop: 225.
FH Key Location/Qualifiers
FH 1..421
FH /organism="Homo sapiens"
FH /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
FH XhoI; Cloned unidirectionally. Primer: Oligo dT. Uninduced
FH exponentially growing neuroepithelial cells
FH (Ntera-2/cl.D1). Average insert size: 1.0 kb; Uni-ZAP XR
FH vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3',
FH adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
FH source
FH 1..421
FH /organism="Homo sapiens"
FH /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
FH XhoI; Cloned unidirectionally. Primer: Oligo dT. Uninduced
```

```
/clone="663740"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
complement(<1..>421)
FT mRNA
SQ Sequence 421 BP; 164 A; 55 C; 67 G; 135 T; 0 other;

Query Match      5.0%; Score 37; DB 82; Length 421;
Best Local Similarity 43.4%; Pred. No. 7.52e-08;
Matches 36; Conservative 16; Mismatches 31; Indels 0; Gaps 0;

Db 308 aaataagagtggcagccataaggaatactatttataaaataacagagttatagagcta 367
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 632 ARATNATNYTGARACNAARNAACNATNATAYAAARYTNAAYGGGTNWSNGARMGNG 691
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 ctttaagaagaatgaactttgg 390
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 692 AYTNAARAARWSNGTYTGTGG 714
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
LOCUS AA227473 421 bp mRNA EST 24-FEB-1997
DEFINITION zrl8all.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
ACCESSION AA227473
NID g1849027
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 421)
AUTHORS Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
Parsons J., Rifkin L., Rohlfing T., Tan F., Trevaskis E.,
Waterston R., Williamson A., Wohlmann P. and Wilson R.
WashU-Merck EST Project
Unpublished (1995)
TITLE Contact: Willson RK
JOURNAL WashU-Merck EST Project
COMMENT Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41ml3 fwd. ET from Amersham
High quality sequence stop: 225.
FEATURES
source
1..421
/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; Cloned unidirectionally. Primer: Oligo dT.
Uninduced, exponentially growing neuroepithelial cells
(Ntera-2/cl.D1). Average insert size: 1.0 kb; Uni-ZAP XR
vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3',
adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
/clone="663740"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
complement(<1..>421)
BASE COUNT      164 a  55 c  67 g  135 t
ORIGIN

Query Match      5.0%; Score 37; DB 63; Length 421;
Best Local Similarity 43.4%; Pred. No. 7.52e-08;
Matches 36; Conservative 16; Mismatches 31; Indels 0; Gaps 0;

Db 308 aaataagagtggcagccataaggaatactatttataaaataacagagttatagagcta 367
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
```


521	CGGACACCAAGAAGTTCCTGTGCTCGCTCTTCGCGCCCCCGCTCGCTCGATGACTAGACG	580
619	agccgctgcgcgcctgcgcgcctcgtgctgcgagcgcgcgaagccgcgcgcctgcgcgcctga	678
581	AGACCATCCAGGCATGCCACTGCCTCGTGCAGAGTGAAGGACCGCTGCGCCCGGTCGA	640
679	tgcgcagtaeaggcttcgcctgc	730
641	TGTCGCGCTTCGGCTTCGCCCTCCCTGCGCCGACATGCTTGATGGACCGTTTCC	692

3	MMU43321	2421 bp	DNA	ROD	24-FEB-1996
LOCUS	Mus musculus	putative	transmembrane receptor	(fizzled 8)	gene
DEFINITION	complete cds.				
ACCESSION	U43321				
NID	91151259				
KEYWORDS	house mouse.				
SOURCE	Mus musculus				
ORGANISM	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Mus.				
REFERENCE	1 (bases 1 to 2421)				
AUTHORS	Wang, Y., Macke, J.P., Abella, B.S., Andreasson, K., Worley, P., Gilbert, D.J., Copeland, N.G., Jenkins, N.A. and Nathans, J.				
TITLE	A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene <i>fizzled</i>				
JOURNAL	J. Biol. Chem. 271 (8), 4468-4476 (1996)				
MEDLINE	96224032				
REFERENCE	2 (bases 1 to 2421)				
AUTHORS	Abella, B., Wang, Y., Macke, J.P. and Nathans, J.				
TITLE	Direct Submission				

```

Direct Submission
Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and
Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street,
Baltimore, MD 21205
FEATURES
Location/Qualifiers
1..2421
/organism="Mus musculus"
/chromosome="18"
/map="between Tpl2 and Cdh2"
188..2245
/gene="frizzled 8"
/note="putative transmembrane receptor"
/codon_start=1
/product="transmembrane receptor"
/db_xref="pid:g1151260"
/translation="MEEGYLLEVTSLLAAVLQSSGDAASAASAKELACQETV
GIGNIYMPQNFHDQDEAGLEVHQFWPLVEIQSPDLKFFLCSTYFLLCLL
PLPSPVSRCEAKAGCAPLMKRYGFAMPDRMKDRLEPQGNPDLCMDYINRTDI
PSPRRLLPPPPPEQGPSSGSHRPPGARGPDRHSGSGDAAAAPGGLLAGFVS
GGGAALPCGQCCAPAMVSVSSSRHPLYNPKVTQIANCALPCHNPFPSODERA
WIGLWSVLCFVSTPATVSTFLIDMEREFKYPRTLFIQSACVLFVSQVLYRLVA
VACSGGAPAGGRCGAGGAGAAAGAGAGAGASSPGARGEYEELGAVQHVRYET
SIVALLVLLVDFVAGSITWVLLSLTWFLAAGMKNGEATAGYSQYFHLAAWL
SIAVLSLVDDQPVACQYVGNOSLDNLRGFLVAPLYLFIQTMLFAGFVS
RSVTKQGGPPTKHLKRLMRLGLFTVLYTPAAVVAACVLFQHNRPWEAAE
CLRDLPQDAARPDPYAFMLKLYVLTISGVVWMSGKLTLESWRLALCTRCCL
AAVAGAGGSGPGSGPGPGGGGGGGGGSLYSDVSTGTLWRSGTASSVSYSP

```

BASE COUNT	403 a	772 c	779 g	467 t	ORIGIN
Query Match	2.8%;	Score 56;	DB 87;	Length 2421;	
Best Local Similarity	59.4%;	Pred. No. 1.17e-17;			
Matches 177;	Conservative 0;	Mismatches 121;	Indels 0;	Gaps	
306	tgccgttgtgcaaaagccatcggttacacabctacacttatatgccaacacagtttcaaccacg	365			
395	tcacgctgtgccacggcatcgatcgataccgacacatccggctctgcccaacctggttgccccacg	454			
366	abacgcaagatgagcgcggtcttagaggttgcaaccagtttggccgctggttgagdgatacacagt	425			

RESULT	LOCUS	DEFINITION	ACCESSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE
4	RATRFRZH	1912-bp Rattus norvegicus Drosophila polarity gene (frizzled) homologue mRNA, complete cds.	U02530	frizzled gene; homologue; polarity gene.	Rattus norvegicus (strain Sprague-Dawley) osteosarcoma cDNA to mRNA.	Rattus norvegicus	1 (bases 1 to 1912)	Chan, S.D.H., Karpf, D.B., Fowlkes, M.E., Hooks, M.S., Vuong, V., Bambino, T., Liu, M.Y.C., Arnaud, C.D., Strewler, G.J. and Nussenson, R.A.	Two homologs of the Drosophila polarity gene frizzled (fz) are widely expressed in mammalian tissues	J. Biol. Chem. 267, 25202-25207 (1992)	93094228

	Query Match	2.2%	Score 45;	DB 90;	Length 1912;
	Best Local Similarity	60.2%;	Pred. No. 1.59e-10;		
Matches	177; Conservative	0;	Mismatches 114;	Indels	3; Gaps 2;
Db	228 gctgtgcacdgacatcgccatacagcaccaaccatcatgcccaacctctctgggcaacagaa	287			
NOY	399 GCVTGCCAGGGCATCGAATACCAGAAGATGCGCTGCCTCAACTGCTGGGCACAGAGAC	458			
	288 ccagaggagdcggggcttgagggttatcaattcttcccgcctggttaagggttcagtgtctc	347			
	CATGAAGGAGGTGCTGAGTAGGCTGGCCCTTGATATCCGCGTGGTATGACACATGCCA	518			

[illegible]

Query Match	2.08;	Score 41;	DB 80;	Length 1923;
Best Local Similarity	59.5%;	Pred. No. 4.63e-08;		
Matches 175;	Conservative 0;	Mismatches 116;	Indels 3;	Gaps 2;
Db	267	gcctgcacggacatcgctctacaacagaccatcatgcccacactctctgaggccacacgaa	326	
QY	399	GCTGTGCCACGGATCGATACAGAACATCGCGCTGCCAACTGCTGGGCCACGAGAC	458	
Db	327	ccaggaggacgcaggcctagaggtgcaccagttctatccgctggtgaagtgcaagtctc	386	
QY	459	CATGAAGGAGGTGCTGAGCAGCGCGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCA	518	
Db	387	gccgaactgcgcttctctctgctccatgtacgcaccgcgtgtgcaccg-tg--ctgga	443	
QY	519	CCGGGACCCAGAAAGTTCCTGTGCTGCTGCTTCCTGGCCCCGCTGCTCGTCACTAG	578	
Db	444	acaggccatcccgcgcgtctctatctgtgagcgcgcgcgcgcgcgcgcgcgcgcgcct	503	

WATERBURY (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelligentGenetics, Inc.

MPsrch_on n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Dec 9 10:16:10 1997; MasPar time 439.22 Seconds
Tabular output not generated. 964.187 Million cell updates/sec

Title: >848439-3-trans
Description: (1-825) from translate.seq
Perfect Score: 688
N.A. Sequence: 1 WSGNCGMNGGNYTNTTYT.....SNATMGNAAATNCARTGY 825
Comp: SWNCGNKCNCNCNANARRA.....WNTANKCNTTYRANGTYACR

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 707517 seqs, 256659390 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99
EST-STS-TWO
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
170:EST170 171:EST171 172:EST172 173:EST173 174:EST174

Statistics: Mean 14.817; Variance 4.177; scale 3.547

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	345	50.1	462 174	W77968	zd70ell.r1 Soares fet	0.00e+00
2	319	46.4	396 178	AA073862	mm97f07.r1 Stratagene	0.00e+00
3	304	44.2	426 172	H87071	y974d07.r1 Homo sapie	0.00e+00
4	279	40.6	406 61	H16121	y128c05.r1 Homo sapie	0.00e+00
5	277	40.3	414 156	AA024771	ze76g10.r1 Soares fet	0.00e+00
6	242	35.2	344 61	H14917	y126d03.r1 Homo sapie	2.91e-283
7	210	30.5	299 148	W08345	mb41f02.r1 Soares mou	3.91e-237
8	163	23.7	231 65	H29095	ym31e03.r1 Homo sapie	2.17e-170
9	151	21.9	250 172	AA061047	mj78a12.r1 Soares mou	1.30e-153
10	63	9.2	250 172	MA9402	CEL08B2 Caenorhabditi	1.52e-37
11	57	8.3	360 141	C09584	C.elegans cdna clone	1.58e-30
12	52	7.6	360 89	CELX117B4F	C.elegans cdna clone	7.36e-25
13	41	6.0	360 89	CELX079D1F	C.elegans cdna clone	3.40e-13
14	40	5.8	256 47	H45312	y65b03.r1 Homo sapie	3.35e-12
15	37	5.4	206 101	N63268	yz88d01.s1 Homo sapie	2.64e-09
16	37	5.4	418 101	N62714	yz76g05.s1 Homo sapie	2.64e-09
17	36	5.2	341 77	CELX119F4R	C.elegans cdna clone	2.27e-08
18	34	4.9	336 123	HSBA08082	H. sapiens partial cd	1.49e-06
19	33	4.8	227 184	AA017479	ze38h10.r1 Soares ret	1.14e-05
20	33	4.8	350 187	AA144540	mr17f12.r1 Soares mou	1.14e-05
21	33	4.8	414 194	AA169546	zo89b08.r1 Stratagene	1.14e-05
22	33	4.8	428 95	N41923	yy07a02.r1 Homo sapie	1.14e-05
23	33	4.8	433 194	AA169521	zo89a08.r1 Stratagene	1.14e-05
24	33	4.8	443 113	W03474	za08a10.r1 Soares mel	1.14e-05
25	33	4.8	458 23	R42865	yy06a10.s1 Homo sapie	1.14e-05
26	32	4.7	219 129	HSC0EA112	H. sapiens partial cd	8.27e-05
27	32	4.7	226 156	AA027879	zk05e04.r1 Soares pre	8.27e-05
28	32	4.7	242 119	W44804	zb98g09.s1 Soares par	8.27e-05
29	32	4.7	233 135	R58334	G2338 Fetal heart Hom	8.27e-05
30	32	4.7	282 81	T35329	EST82912 Homo sapiens	8.27e-05
31	32	4.7	384 95	N42436	yy07e11.r1 Homo sapie	8.27e-05
32	32	4.7	473 127	W90442	zh78c02.s1 Soares fet	8.27e-05
33	32	4.7	584 4	T67709	yz27c11.s1 Homo sapie	8.27e-05
34	32	4.7	522 109	HUM069H03B	Human fetal brain cdn	5.73e-04
35	31	4.5	252 48	HUM055B07B	Human fetal brain cdn	5.73e-04
36	31	4.5	252 109	HUM069D03B	Human fetal brain cdn	5.73e-04
37	31	4.5	330 13	T01376	W8202097 Caenorhabdi	5.73e-04
38	31	4.5	364 42	R27662	yh64c07.r1 Homo sapie	5.73e-04
39	31	4.5	390 1	CELX025G2R	C.elegans cdna clone	5.73e-04
40	31	4.5	404 13	T01720	W8202441 Caenorhabdi	5.73e-04
41	31	4.5	429 8	T85962	y626b09.r1 Homo sapie	5.73e-04
42	31	4.5	443 144	N94473	zb80a06.s1 Soares sen	5.73e-04
43	31	4.5	574 135	N98880	zb87h03.s1 Homo sapie	5.73e-04
44	30	4.4	344 14	T09517	0113m3 Plasmodium fal	3.77e-03

ALIGNMENTS

1 W77968 462 bp mRNA EST 17-OCT-1996
zd70ell.r1 Soares fetal heart NbHH19W Homo sapiens cdna clone
LOCUS 346028 5'
DEFINITION W77968
ACCESSION 91388502
NID
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 462)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevisan,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1550 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 424.
Location/Qualifiers
1..462
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7m3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCACTGCTGAAGTGGCGCCGACATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7m3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W."
/clone="346028"
/clone_lib="Soares fetal heart NBHL19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1..>462
BASE COUNT 128 a 133 c 106 g 91 t 4 others
ORIGIN
Query Match 50.1%; Score 345; DB 174; Length 462;
Best Local Similarity 58.0%; Pred. No. 0.00e+00;
Matches 266; Conservative 108; Mismatches 81; Indels 4; Gaps 4;
Db 6 tcgtctctgcgcccgctgtgctcgatgacccatagacgagaccatccagccatgcacctcg 65
: : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
QY 229 WSNNTTYGCCNCNGTNTGTYTNGAYGAYTNGAYGARACATNATNCARCCNTGYCAWSN 288
: : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Db 66 ctctgn-tgcagtggaaggacgctgcgcccggtcatgtccgcttcgn-ttcccctgg 123
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
QY 289 YTNTGYTNCARGTNAARGATMGNTGYCNCNGTNTGWSNGCNTTYGGNTTYCCNTGG 348
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Db 124 cccgacatgcttgagtcgcagcgtttcccccaggagacacaccttgcacccctccgc 183
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
QY 349 CCNGAYATGTYNGARTGVGYMGNTTYCCNCARGAAYAYTNGYATNCCYNTNGCN 408
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Db 184 tagcagaccacctctctgccgccacgaggaagctccaaaggtatgtgaagcctgc 243
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
QY 409 WSNWSNGAYCAYTYNTNCCNGCNCACGARGCNCNCNAARGTNTGYGRCNTGYAAR 468
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
Db 244 aataaaatgatgatgacacagacataatgaaacgcttttaaaaaatgatttgcactg 303
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :
QY 469 AYAARAAYGAYGAYAAAYGAYATNATGARACNTNTGYAARAAYGAYTTGCGNTN 528
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| :

Db	304	aaataaaagtgaggagataaacctacatcaacccgagataccacaaaatcatcctctggaagac	363
QY	529	AARATNAAGTNAARGARATNACNTAYATNAAATMGNGAYACNAARATNATNYNGAR-AC	587
Db	364	caagagcaagaccatttacaagctgaacgggtgtgtccgaaagaggacacctgaagaaatcgg	423
QY	588	NAARWSNAARACATNATYAAARTNAAATYAGNGTNGWSNARMGNGAYTIN-AARAARWSNG	646
Db	424	tgtctgtcctcaagacagcttcagtcgacgtgcacctgtgaag	462
QY	647	TNYTNTGGYNAARGAYWSNYTNCARTGYACNTGYGARG	685
RESULT	2		
LOCUS	AA073862	396 bp	EST
DEFINITION	mm97f07_r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone		04-OCT-1996
	536389 5'	similar to TR:G1151260	G1151260 TRANSMEMBRANE RECEPTOR.
	;		
ACCESSION	AA073862		
NID	g1595602		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 396)		
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geiseli,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.		
TITLE	The WashU-HMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of MedicineP 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:323325 Seq primer: primer name ambiguous High quality sequence stop: 390. Location/Qualifiers 1..396 /organism="Mus musculus" /strain="NIH/Swiss" /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; cloned unidirectionally. Primer: Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTATTTTATTTT 3" /clone="536389" /clone_lib="Stratagene mouse heart (#937316)" /sex="pooled" /dev_stage="13 day embryos" /lab_host="SOLR (kanamycin resistant)" <1..396		
FEATURES	source		
	BASE COUNT	93 a 126 c 103 g 74 t	
	ORIGIN		
	Query Match	46.4%	Score 319; DB 178; Length 396;
	Best Local Similarity	61.7%	Pred. No. 0.00e+00;
	Matches	245; Conservative	88; Mismatches 62; Indels 2; Gaps 2;
Db	1	agcaggcggcgc-tggattccctgtgctatgaagcagtcgaccgcggacaccacagaagt	59
QY	161	ARCARGCNGCNGTGGATNCCNTNGTATGAARCATGTCYCAACGAYACNAARAAT	220

187 000000 187

181 gcaatgc
 ||:|:|:

M O S E R

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Dec 9 10:39:19 1997; MasPar time 224.95 Seconds
Tabular output not generated. 1007.870 Million cell updates/sec

Title: >848439-3-trans
Description: (1-825) from translate.seq
Perfect Score: 688
N.A. Sequence: 1 WSGCGNMGGNGYNTNTYTT.....SNATNMGAARYTNCARTGY 825
Comp: SWNGCKNCCNRRANAARRA.....WNTANKCNTYRANGYACR

Scoring table: TABLE default
Gap 6
Nmatch STD : Dbase 0; Query 0
Searched: 359085 seqs, 137405154 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS-THREE
1: EST199 2: EST200 3: EST201 4: EST202 5: EST203 6: EST204
7: EST205 8: EST206 9: EST207 10: EST208 11: EST209 12: EST210
13: EST211 14: EST212 15: EST213 16: EST214 17: EST215
18: EST216 19: EST217 20: EST218 21: EST219 22: EST220
23: EST221 24: EST222 25: EST223 26: EST224 27: EST225
28: EST226 29: EST227 30: EST228 31: EST229 32: EST230
33: EST231 34: EST232 35: EST233 36: STS1 37: STS2 38: STS3
39: STS4 40: STS5 41: STS6 42: STS7 43: STS8 44: STS9 45: STS10
46: STS11 47: STS12 48: STS13
EST-STS-FOUR
49: gNEST1 50: gNEST2 51: gNEST3 52: gNEST4 53: gNEST5
54: gNEST6 55: gNEST7 56: gNEST8 57: gNEST9 58: gNEST10
59: gNEST11 60: gNEST12 61: gNEST13 62: gNEST14 63: gNEST15
64: gNEST16 65: gNEST17 66: gNEST18 67: gNEST19 68: gNEST20
69: gNEST21 70: gNEST22 71: gNEST23 72: gNEST24 73: gNEST25
74: gNST5 75: gNEST1 76: gNEST2 77: gNEST3 78: gNEST4
79: gNEST5 80: gNEST6 81: gNEST7 82: gNEST8 83: gNEST9
84: gNEST10 85: gNEST11 86: gNEST12 87: gNEST13 88: gNEST14
89: gNEST15 90: gNEST16 91: gNEST17 92: gNEST18 93: gNEST19
94: gNEST20 95: gNEST21 96: gNEST22 97: gNEST23 98: gNEST24
99: gNEST25 100: gNEST26 101: gNEST27 102: gNEST28
103: gNEST29 104: gNEST30 105: gNEST31 106: gNST5 107: ueEST1
108: ueEST2

Database: EST-STS-FOUR
49: gNEST1 50: gNEST2 51: gNEST3 52: gNEST4 53: gNEST5
54: gNEST6 55: gNEST7 56: gNEST8 57: gNEST9 58: gNEST10
59: gNEST11 60: gNEST12 61: gNEST13 62: gNEST14 63: gNEST15
64: gNEST16 65: gNEST17 66: gNEST18 67: gNEST19 68: gNEST20
69: gNEST21 70: gNEST22 71: gNEST23 72: gNEST24 73: gNEST25
74: gNST5 75: gNEST1 76: gNEST2 77: gNEST3 78: gNEST4
79: gNEST5 80: gNEST6 81: gNEST7 82: gNEST8 83: gNEST9
84: gNEST10 85: gNEST11 86: gNEST12 87: gNEST13 88: gNEST14
89: gNEST15 90: gNEST16 91: gNEST17 92: gNEST18 93: gNEST19
94: gNEST20 95: gNEST21 96: gNEST22 97: gNEST23 98: gNEST24
99: gNEST25 100: gNEST26 101: gNEST27 102: gNEST28
103: gNEST29 104: gNEST30 105: gNEST31 106: gNST5 107: ueEST1
108: ueEST2

Statistics: Mean 14.890; Variance 4.327; scale 3.441
pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result Query
SUMMARIES

No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	357	51.9	458	33	AA105749	ml84a03.r1 Stratagene	0.00e+00
2	319	46.4	396	54	AA073862	mm97f07.r1 Stratagene	0.00e+00
3	235	34.2	394	101	MMAA60087	va52g08.r1 Soares mou	7.48e-265
4	210	30.5	299	94	MM3451	mb41f02.r1 Soares mou	6.10e-230
5	148	21.5	184	20	AA209833	mo90h01.r1 Beddington	6.84e-145
6	145	21.1	523	56	AA128822	mr03c11.r1 Soares mou	7.24e-141
7	145	21.1	523	6	AA128822	mr03c11.r1 Soares mou	7.24e-141
8	136	19.8	287	32	AA082155	ze88f06.r1 Soares fet	8.23e-129
9	55	8.0	502	21	AA020088	mh49d08.r1 Soares mou	1.99e-27
10	37	5.5	213	28	AA023595	mh80d03.r1 Soares mou	5.16e-10
11	38	5.4	421	82	HS1147060	zr18a11.s1 Stratagene	4.29e-09
12	37	5.4	421	63	AA227473	zr18a11.s1 Stratagene	4.29e-09
13	36	5.2	512	29	AA044908	zf51e10.r1 Soares ret	3.44e-08
14	35	5.1	432	29	AA030216	mh87g08.r1 Soares mou	2.66e-07
15	35	5.1	519	39	G08452	human STS CHLC.ATA11D	2.66e-07
16	34	4.9	288	36	DM57F11T	D. melanogaster STS d	1.97e-06
17	34	4.9	496	10	AA013810	mh06e01.r1 Soares mou	1.97e-06
18	33	4.8	241	21	AA019228	ze59c03.r1 Soares ret	1.41e-05
19	33	4.8	252	37	G03067	human STS WI-5108	1.41e-05
20	33	4.8	350	56	AA144540	mr17f12.r1 Soares mou	1.41e-05
21	31	4.5	407	89	HSAA62672	zs23f08.r1 Soares NDR	6.23e-04
22	31	4.5	451	39	G09558	human STS CHLC.GCT8C0	6.23e-04
23	30	4.4	177	36	DM52H6T	D. melanogaster STS d	3.85e-03
24	30	4.4	199	10	AA017992	mh45a07.r1 Soares mou	3.85e-03
25	30	4.4	213	10	AA018016	mh45c07.r1 Soares mou	3.85e-03
26	30	4.4	305	21	AA019656	ze61f11.s1 Soares ret	3.85e-03
27	30	4.4	339	43	G26024	human STS EST186890	3.85e-03
28	30	4.4	402	1	N97824	1353C3 czapppfD2.1, D	3.85e-03
29	30	4.4	425	43	G24362	ze71h06.s1 Soares fet	3.85e-03
30	30	4.4	425	43	G24362	human STS WI-13950	3.85e-03
31	30	4.4	435	34	AA141629	CK02453.3prine Drosop	3.85e-03
32	30	4.4	440	26	N54746	yz30a03.s1 Soares mul	3.85e-03
33	30	4.4	455	80	CPAAJ4647	CpEST.119 psKriminusc	3.85e-03
34	30	4.4	455	62	AA224647	CpEST.119 psKriminusc	3.85e-03
35	30	4.4	471	44	G27881	human STS SHGC-34074	3.85e-03
36	30	4.4	503	92	NM1157919	mv26g08.r1 GuayWoodfo	3.85e-03
37	30	4.4	503	70	AA241149	mv26g08.r1 GuayWoodfo	3.85e-03
38	29	4.2	170	36	DM155E2T	D. melanogaster STS d	2.26e-02
39	29	4.2	298	1	N97343	0109M3 gmbpffH3.1, G	2.26e-02
40	29	4.2	367	75	AT15723	22538 lambdaB-PRL2 Ara	2.26e-02
41	29	4.2	386	56	AA163930	mr19c02.r1 Soares mou	2.26e-02
42	29	4.2	427	39	G09767	human STS CHLC.ATA15C	2.26e-02
43	29	4.2	449	66	AA237923	mx14f07.r1 Soares mou	2.26e-02
44	29	4.2	491	101	MMAA63807	mr18D06.r1 Soares mou	2.26e-02
45	29	4.2	522	73	C22763	Dictyostellium discoi	2.26e-02

ALIGNMENTS

RESULT	1
LOCUS	AA105749 458 bp mRNA
DEFINITION	ml84a03.r1 Stratagene mouse kidney (#937315) Mus musculus CDNA clone 518668 5', similar to TR:G1151260 G1151260 TRANSMEMBRANE RECEPTOR. ;
ACCESSION	AA105749
NID	g1654838
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 458) Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE	The WashU-HMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	


```

QY 461 CNTGYAARAAYARAAYGAYGAYAAAYGAYATNATGGARACNYTNTGYAARAAYGAYT 520
Db 359 tcgcactgaaatcaaaagtgagagataacgtacat 395
QY 521 TYGCNTNARATNARGTNAARGATNACNTAYAT 557

RESULT 3
ID MAA60087 standard; RNA; EST; 394 BP.
AC AA260087;
NI g1896590;
DT 19-MAR-1997 (Rel. 51, Created)
DE va52q08.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 735038 5'.
KW EST.
OS Mus musculus (house mouse)
OC Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Mus.
RN [1]
RP 1-394
RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA Waterston R.;
RA "The WashU-HHMI Mouse EST Project";
RL Unpublished.
CC Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC Washington University School of MedicineP 4444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: mouseest@wustl.edu This clone is available
CC royalty-free through LNL : contact the IMAGE Consortium
CC (info@image.llnl.gov) for further information. MGI:452086 Seq
CC primer: -28ml3 rev2 Et from Amersham.
FH Location/Qualifiers
FH
FH source
FH 1..394
FH /organism="Mus musculus"
FH /strain="C57BL/6J"
FH /note="Vector: pT73D-Pac (Pharmacia) with a modified
FH polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FH was primed with a Not I - oligo(dT) primer [5',
FH TGTTACCAATCTGAAGTGGAGCGCGCTTATTTTTTTTTTTT 3'], o
FH
FH total mouse RNA [provided by Minoru Ko, Wayne State Univ.];
FH double-stranded cDNA was ligated to Eco RI adaptors
FH (Pharmacia), digested with Not I and cloned into the Not I
FH and Eco RI sites of the modified pT7T3 vector. Library wen
FH
FH through one round of normalization, and was constructed by
FH Bento Soares and M. Fatima Bonaldo."
FH /clone="735038"
FH /clone_lib="Soares mouse 3NME12 5"
FH /sex="unknown"
FH /dev_stage="12.5dpc total fetus"
FH /lab_host="DH10B"
FH mRNA
FH <1..>394
SQ Sequence 394 BP; 104 A; 96 C; 118 G; 76 T; 0 other;

Query Match 34.2%; Score 235; DB 101; Length 394;
Best Local Similarity 57.5%; Pred. No. 7.48e-265;
Matches 164; Conservative 71; Mismatches 50; Indels 0; Gaps 0;

Db 1 aaggagataacgtacatcaaacagacacacatctctctggagacaagagacacc 60
QY 541 AARGARATNACNTAYATNAYMGAYACNAARATNTNGARACNAARWSNAARACN 600
Db 61 attacaagctgaacggcgctgcgaagggaccctgaagaaatccgtgctggtcctaaa 120
QY 601 ATNTAAYARTNAYGGNGTWSNGARMGNGAYYTNAARAARWSNGTNTTGCTTNAAR 660
Db 121 gagagctgcagtcacgtctgaggagatgaacgacataacgcctcgtatctggtcatg 180

```

```

QY 661 GAYWSNTNCTAGTGYACNTGYGARGATGAAYGAYATNAAAYCNCNTAYITNGTATG 720
Db 181 ggacagaagcagggcgagctggtgatcacctccctgaaacggtggcagaaggccag 240
QY 721 GGCARAAARCAGGGGNGGARYTNGTNTATNACNWSNTNAARMGNTGGCARAAGNCAR 780
Db 241 agagagttcaagcgcatctccgcagcatccgcgaagctgcaatgc 285
QY 781 MGNGARTTYAARMGNATNWSMGNWSNATNMGNAARYTNCARTGY 825

RESULT 4
ID MM3451 standard; RNA; EST; 299 BP.
AC W08345;
NI g1282366
DT 27-APR-1996 (Rel. 47, Created)
DE 05-MAR-1997 (Rel. 51, Last updated, Version 2)
DE mb411f02.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 331995
DE 5'.
KW EST.
OS Mus musculus (house mouse)
OC Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Mus.
RN [1]
RP 1-299
RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA Waterston R.;
RA "The WashU-HHMI Mouse EST Project";
RL Unpublished
CC Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC Washington University School of MedicineP 4444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: mouseest@wustl.edu This clone is available
CC royalty-free through LNL : contact the IMAGE Consortium
CC (info@image.llnl.gov) for further information. MGI:213395 Seq
CC primer: mob.REGA+ET High quality sequence stop: 280.
FH Location/Qualifiers
FH
FH source
FH 1..299
FH /organism="Mus musculus"
FH /note="Vector: pT73D (Pharmacia) with a modified
FH polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FH was primed with a Not I - oligo(dT) primer [5',
FH TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3'],
FH double-stranded cDNA was size selected, ligated to Eco RI
FH adaptors (Pharmacia), digested with Not I and cloned into
FH the Not I and Eco RI sites of a modified pT7T3 vector
FH (Pharmacia). Library went through one round of
FH normalization to a Cot = 5. Library constructed by Bento
FH Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr
FH
FH Minoru Ko (Wayne State University)."
FH /clone="331995"
FH /clone_lib="Soares mouse p3NMF19.5"
FH /dev_stage="19.5 dpc total fetus"
FH /lab_host="DH10B (ampicillin resistant)"
FH mRNA
FH <1..>299
SQ Sequence 299 BP; 79 A; 95 C; 72 G; 53 T; 0 other;

Query Match 30.5%; Score 210; DB 94; Length 299;
Best Local Similarity 59.5%; Pred. No. 6.10e-230;
Matches 175; Conservative 68; Mismatches 46; Indels 5; Gaps 5;

Db 11 gctcgtcactcgtctcgtgcag-tgaaggaccctgagc-cgcgtcatctgcgcctt 67
QY 276 RCCNTGTCAYWSNTNTGTYGTNCARGTNAARGAYMGNTGTCNCNGTNGTNGCNCNT 335
Db 68 cg-cttccctcgtgcacacatcgtcgtgagtcgcagcgtttcccgag-acaacgacctcg 125

```

```

QY 336 YGGNTTYCCNTGGCNGAYATGYTNGARTGYGAYMGNTTYCCNCARGAAYAYGAYTYNTG 395
      :| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 126 catccctcctagtagcgaccacctcctcccgccacaggaagctcccaagtggtg 185
      :|| || || || || || || || || || || || || || || || || || || || || ||
QY 396 YATWCCNTYTCNNWSNWSNGAYCAYTYTNTCCNGCNCAGARGCNCNAARGTNTG 455
      :|| || || || || || || || || || || || || || || || || || || || || ||
Db 186 tgaagcctgcaaaacaaagaatgagcagcaacacacatcatcggaaccccttgtaaaaa 245
      :|| || || || || || || || || || || || || || || || || || || || || ||
QY 456 YGAGCWTGYAARAAYARAAYGAYGAYAYATNATGGARACNYNTNGYAARA 515
      :|| || || || || || || || || || || || || || || || || || || || || ||
Db 246 tgacttcgactgaaatacaagtgaagagataacgctacatcaacagagacac 299
      :|| || || || || || || || || || || || || || || || || || || || || ||
QY 516 YGAYTTYGCNTNAARATNAARGTNAARGATNACNTAYATNAAYMNGAYAC 569
      :|| || || || || || || || || || || || || || || || || || || || || ||

RESULT 5
LOCUS AA209833 184 bp mRNA EST 29-JAN-1997
DEFINITION m090h01.r1 Beddington mouse embryonic region Mus musculus cDNA
clone 560881 5'.
ACCESSION AA209833
NID G1808152
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 184)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:341673
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 147.
FEATURES
Location/Qualifiers
source 1..184
/organism="Mus musculus"
/strain="C57BL6 x DBA"
/notes="vector: pCMV-SPORT; Site_1: SalI; Site_2: NotI;
Cloned unidirectionally. Primer: Oligo dT. Gastrulating
embryos were collected at 7.5dpc from C57BL6 x DBA
matings, excluding embryos that had developed head folds
and all extraembryonic tissues. Average insert size: 1.3
kb (range: 0.5 - 3.0 kb). Referenced in Development 121,
2479-2489 (1995)"
/clone="560881"
/clone_lib="Beddington mouse embryonic region"
/sex="pooled"
/dev_stage="7.5dpc"
/lab_host="DH12S"
<1..>184

mRNA 61 a 45 c 48 g 30 t
BASE COUNT
ORIGIN
Query Match 21.5%; Score 148; DB 20; Length 184;
Best Local Similarity 58.9%; Pred. No. 6.64e-145;
Matches 109; Conservative 45; Mismatches 30; Indels 1; Gaps 1;

```

```

Db 1 gatacgtacatcaacagagacacacacatcatcctggagacaaagacagaccattta 60
      :|| || || || || || || || || || || || || || || || || || || || || ||
QY 546 RATNACNTAYATNAAYMNGAYACNAARATNATNTNGARACNAARWSNAARACNATNTA 605
      :|| || || || || || || || || || || || || || || || || || || || || ||
Db 61 caagctgaacggcgtgtccgaag-gacctgaagaaatccgtgctggtcctcaagacag 119
      :|| || || || || || || || || || || || || || || || || || || || || ||
QY 606 YAARYTNAAYGGNGTWSNGARMGNGAYTTNAARAARWSNGTNTTGGYTNAARGAYS 665
      :|| || || || || || || || || || || || || || || || || || || || || ||
Db 120 cctgcagtcacctgtgagagatgaacgacatcaacgctcgtatcgtgctatgggaca 179
      :|| || || || || || || || || || || || || || || || || || || || || ||
QY 666 NYTNCARTGYACNTGYGARGARATGAAYATNAAYGCNCCNTAYTYNGTATGGNCA 725
      :|| || || || || || || || || || || || || || || || || || || || || ||
Db 180 gaagc 184
      :|| ||
QY 726 RAARC 730

RESULT 6
LOCUS AA128222 523 bp mRNA EST 16-FEB-1997
DEFINITION mr03c11.r1 Soares mouse 3NBMS Mus musculus cDNA clone 596372 5'.
ACCESSION AA128222
NID G1681779
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 523)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:361804
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 469.
FEATURES
Location/Qualifiers
source 1..523
/organism="Mus musculus"
/strain="C57BL/6J"
/notes="vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTATCAATCTGAAGTGGAGCGCGCGCTGTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M.Fatima Bonaldo."
/clone="596372"
/clone_lib="Soares mouse 3NBMS"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"
<1..>523

mRNA 126 a 142 c 129 g 126 t
BASE COUNT
ORIGIN

```


/lab_host="DH10B (ampicillin resistant)"
<1..>287

BASE COUNT 62 a 84 c 85 g 51 t 5 others
ORIGIN

Query Match 19.8%; Score 136; DB 32; Length 287;
Best Local Similarity 58.6%; Pred. No. 8.23e-129;
Matches 102; Conservative 41; Mismatches 30; Indels 1; Gaps 1;

Db 1 ggcacagacagcttgagtcgaccttgaggagatgaacacacacacgcgcctatc 60
Qy 653 GGYNAARGAYNSYNTNCARTGYACNTGYGARGARATGAAYATNAAGNCNCNTAYY 712
Db 61 tggctcatggcagacaaacaggggtggagctggtgatcacctcggtgaagcggtggcaga 120
Qy 713 TNGINATGGGNCARACARGGNGGARYTNGTNGTATNACNSGNTNARMGNTGGCARA 772
Db 121 agggcgagagagtgctcaagcgatctcccggnngnncatccgcaagctcgatgyc 174
Qy 773 ARGNCARMGNGARTTYAARMGNATNWSMGNW-SNATMNGNAARYTNCARTGY 825

RESULT 9 AA020088 502 bp mRNA EST 21-JAN-1997
LOCUS mh49d08.r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
DEFINITION clone 445839 5' similar to PIR:A45054 A45054 probable intercellular
signal transducer or transmitter Fz-1 ;

ACCESSION AA020088
NID g1483933
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 502)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:271175

Possible reversed clone: similarity on wrong strand
Seq primer: -28M13 rev2 from Amer-sham
High quality sequence stop: 444.

FEATURES source
Location/Qualifiers
1..502
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAACTGAGTGGGAGCGCGCGGAAATTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
/clone="445839"
/clone_lib="Soares mouse placenta 4NDMP13.5 14.5"
/sex="unknown"

/dev_stage="adult"
/lab_host="DH10B"
<1..>502

BASE COUNT 77 a 143 c 179 g 102 t 1 others
ORIGIN

Query Match 8.0%; Score 55; DB 21; Length 502;
Best Local Similarity 39.8%; Pred. No. 1.99e-27;
Matches 119; Conservative 44; Mismatches 132; Indels 4; Gaps 4;

Db 99 ggggaactctgcacattgagtgctgtgacctgaagcgaactgttctatgagcgcc 158
Cp 377 GGRANCRKTRTCAYTCNARCATRTCNGCCANGGAAANCCRAANGCNWSCATNACNGN 318
Db 159 tcgcagcctcgcgctgcgcacagcgagcgagcgagcgagcgagcgagcgagcgag 218
Cp 317 GRCANCRKTRTCAYTCNACRACANARNSRT-GRCANGGYTGATNGTTCRNCAR 259
Db 219 -t-a-cggtgcacacagcgctgacatgagcacaggaagaactgagctcgcgctaga 275
Cp 258 RTCRTCNARRCANACNGGNGCRAANARNSRRCANARRAAYTTTNGTTCNGGRTGRCA 199
Db 276 ctgcaccttcacacagggtagaactggtgcacctccagacggcgctcctctgattcgt 335
Cp 198 YTGYYTCATNACNARNGGNATCCANGCNCNGCYGTGCNARNACYTCYTTTCATNGTYTC 139
Db 336 gtggcccgacaggttgccatgatggtctggttacgcgatgctcgtgcacagcgaggga 394
Cp 138 RIGNCNANARRTTNGGNARCKCATRTTYTGRTAYTCNATNCCTGRCANARITGNA 80

RESULT 10 AA023595 213 bp mRNA EST 21-JAN-1997
LOCUS mh80d03.r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA
DEFINITION clone 457253 5' similar to PIR:A45054 A45054 probable intercellular
signal transducer or transmitter Fz-1 ;

ACCESSION AA023595
NID g1487495
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 213)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:274141

Possible reversed clone: similarity on wrong strand
Seq primer: -28M13 rev2 from Amer-sham.
Location/Qualifiers
1..213
/organism="Mus musculus"
/strain="C57BL/6J"

FEATURES source
Location/Qualifiers
1..213
/note="Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAACTGAGTGGGAGCGCGCGGAAATTTTTTTTTTTTTTTT T 3']

572 ARACNATNTNGARACWAAEWSNPARACNATNTAYARYTNAAYGGNGTINSGARMGNG 631
QY

368 ctttaagcaagcaatgaacttttg 390
Db ::|||:::|::|::|

632 APTTNPRAAAEWSNGNTINYITIGS 654
QY

RESULT	13
LOCUS	AJ0044908 512 bp mRNA EST 02-FEB-1997
DEFINITION	Zf5c10.r1 Soares retina N264HR Homo sapiens cDNA clone 380490.5' similar to contains L1.t2 L1 repetitive element ;
ACCESSION	AJ0044908
NID	G1523130
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (Cases 1 to 512)
REFERENCE	Authors Haller,T., Clark,N., Dubouque,T., Elliston,K., Hawkins,M., Holzer,M., Hultman,M., Kucapaa,I., Le,M., Lennon,G., Narra,M., Parsons,J., Ricklin,L., Rohlfing,T., Tan,F., Trevisakis,E., Waterson,R., Williamson,A., Woldmann,P. and Wilson,R. WashU-Merck EST Project Unpublished (1995)
TITLE	JOURNAL

Contact: Wilson RK
Washo-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 285 1800
Fax: 314 285 1810
Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 887 Std Error: 0.00
Seq Primer: -28M13 rev2 from Amersham
High quality sequence stop: 314.

```

FEATURES
source
1.1.512
location/Qualifiers
/organism="Homo sapiens"
/notes="Organ: eye; Vector: pT73D (Pharmacia) with a
modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dR) primer [5',
TGTACCAATCTCAAGTGGAGGCGCGCGCTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). The retinas were obtained from a 55 year old
caucasian and total cellular poly(A)+ RNA was extracted 6
hrs after their removal. The retina RNA was kindly
provided by Roderick R. McInnes M.D. Ph.D. from the
University of Toronto. Library constructed by Bento
Soares and M.Fatima Bonalco."
/clone="380490"
/clone_lib="Soares retina N2b4HR"
/sex="male"
/tissue_type="retina"
/dev_stage="55 year old"
/lab_host="DH10B (ampicillin resistant)"
<1..>512

```

```

Query Match      5.2%   Score 36;   DB 29;   Length 512;
Best Local Similarity 37.9%
Pred. No. 3.44e-08;
Matches 47;   Conservative 20;   Mismatches 57;   Indels 0;   Gaps 0;

Db 338   taaaaattacctcaaatatgatccacacataaaagttaaaactataaaactctagaag 337
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 527   TAAAGATNAAGGTAAAGTAAAGATACNTATTAATAAATGNGAYACNAARATATNTNGARA 586
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 398   aaaaatggtaagagaatgtctctagtactactagagttagggcnaaaatttttttagtggggg 457

```

QY 587 CNAARWNAARACNATNTAYAARYTNAAYGNGTINWSNGARGMGNGAYTTNAARARWSNG 645

Db 458 tatt 461

QV 647 TNYT 650

RESULT	14
LOCUS	AA030216 432 bp mRNA EST 21-JAN-1997
DEFINITION	m87g08.r1 Soares mouse placenta 4NbMP13.14.5 Mus musculus cDNA clone 457982 5' similar to PIR:A45054 A45054 probable intercellular signal transducer or transmitter Fz-1 ;

ACCESSION AA030216
NID g1497354
KEYWORDS EST.
SOURCE house mouse.

ORGANISM *Mus musculus*
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.

1 (bases 1 to 432)
Marra, M., Hallier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Gaisell, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Teising, B., Wylie, T., Lennon, G., Soares, B., Willson, R. and
Waterston, R.
The WashU-HMHI Mouse EST Project
Unpublished (1996)
JOURNAL

Contact: Narra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
TMAGE Consortium (info@image.llnl.gov) for further information.

MG1:274870

Possible reversed clone: similarity o

Seq primer: -28M13 rev2 from Ame

FEATURES
source

source

```
/organism="Mus musculus"
```

```
/strain="C57BL/6J"
```

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand

was primed with a Not I - oligo(dT) primer [5'

[illegible]

constructed by Bento Soares and M. Fatima Bonatdu.
(a) 200-7457002"

```
/clone="45/982"  
/clone lib="goatres mouse n[acent]a 4NbMP13 5 14 5"
```

```
/C:\WINDOWS\system32\cmd.exe /c
```

```
/sex="male"
/development="adult"
```

```
/dev_stage= adult
/lab_host="DH10B"
```

```

/ tab_most = diff
<1.>432

```

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
75 a	137 c	153 a	67 t																																																																																																												

ORIGIN

NOTES

Query Match 5.18; Score 35; DB 29; Length 432;

Best Local Similarity 38.8%; Pred. No. 2.66e-07;
Matches 66; Conservative 25; Mismatches 78; Indels 1; Caps 1;

[illegible]

Db 50 gcacacagggcggtacatggagcacaggaacttgagctggcg-gacactgcacctt 108

[illegible]

cp 249 RCANACNGGCGCRAAANARNWSRCANARRAAATTATTGTTCTCNGGRTGRCATTGTTTCAT 190

W P E R L H (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 07:57:40 1997; MasPar time 1674.10 Seconds
Tabular output not generated. 1329.789 Million cell updates/sec

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCTTCATGGCTT.....AAAAAAGGCGGCGC 2027
Comp: CTTAAGCGGGAAGTACCGGA.....TTTTTTTTTTCGCGCGC

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 362067 seqs, 549138275 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-new3
1: BCT 2: FUN 3: GEN1 4: GEN2 5: HTG1 6: HTG2 7: HUM 8: INV
9: ORG 10: MAM 11: VRT 12: PLN 13: PRO 14: ROD 15: SYN 16: UNC
17: VLR
genbank99

Database: 18: BCT1 19: BCT2 20: BCT3 21: BCT4 22: BCT5 23: BCT6 24: BCT7
25: BCT8 26: BCT9 27: BCT10 28: BCT11 29: GEN1 30: GEN2
31: GEN3 32: HTG1 33: HTG2 34: HTG3 35: INV1 36: INV2 37: INV3
38: INV4 39: INV5 40: INV6 41: INV7 42: INV8 43: INV9 44: INV10
45: INV11 46: MAM1 47: MAM2 48: MAM3 49: VRT1 50: VRT2 51: VRT3
52: VRT4 53: PAT1 54: PAT2 55: PAT3 56: PAT4 57: PAT5 58: PHG
59: PLN1 60: PLN2 61: PLN3 62: PLN4 63: PLN5 64: PLN6 65: PLN7
66: PLN8 67: PLN9 68: PLN10 69: PLN11 70: PRI1 71: PRI2
72: PRI3 73: PRI4 74: PRI5 75: PRI6 76: PRI7 77: PRI8 78: PRI9
79: PRI10 80: PRI11 81: PRI12 82: PRI13 83: PRI14 84: PRI15
85: ROD1 86: ROD2 87: ROD3 88: ROD4 89: ROD5 90: ROD6 91: ROD7
92: ROD8 93: STR 94: SYN 95: UNA 96: VRL1 97: VRL2 98: VRL3
99: VRL4 100: VRL5 101: VRL6 102: VRL7 103: VRL8 104: VRL9
105: VRL10
genbank-new3

Database: 106: BCT 107: GEN1 108: GEN2 109: HTG1 110: HTG2 111: INV
112: MAM 113: VRT 114: PHG 115: PLN 116: PRI1 117: PRI2
118: ROD 119: SYN 120: UNA 121: VRL
Database: u-emb150_99
122: Part1

Statistics: Mean 12.293; Variance 6.329; scale 1.942

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

*

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	959	47.3	1776	85	D50462	Mouse SDF5 mRNA, comp 0.00e+00
2	58	2.9	2334	77	HSU43318	Human putative trans 5.31e-19
3	56	2.8	2421	87	MMU43321	Mus musculus putative 1.17e-17
4	45	2.2	1912	90	RATFR2H	Rattus norvegicus Dro 1.59e-10
5	41	2.0	1923	80	HUMFRIZ	Human frizzled gene p 4.63e-08
6	39	1.9	215	57	I28278	Sequence 5 from paten 7.33e-07
7	38	1.9	4540	90	RATFR22H	Rattus norvegicus Dro 2.85e-06
8	31	1.5	215	57	I28278	Sequence 5 from paten 2.40e-02
9	30	1.5	354	111	OFU89259	Oxytricha fallax 57KD 8.07e-02
10	30	1.5	354	8	OFU89259	Oxytricha fallax 57KD 8.07e-02
11	30	1.5	4397	59	ATHSC701	A.thaliana hsc70-1 ge 8.07e-02
12	31	1.5	38262	37	CELK02A2	Caenorhabditis elegans 2.40e-02
13	28	1.4	354	8	OFU89259	Oxytricha fallax 57KD 8.46e-01
14	28	1.4	354	111	OFU89259	Oxytricha fallax 57KD 8.46e-01
15	28	1.4	565	73	HSANX5511	Human annexin V (ANX5 8.46e-01
16	28	1.4	516	73	HSANX5V11	Human annexin V (ANX5 8.46e-01
17	29	1.4	1073	59	ATCOR47	A.thaliana cor47 mRNA 2.65e-01
18	28	1.4	1444	79	HUMATC	Human placental antic 8.46e-01
19	28	1.4	1454	55	I08832	Sequence 3 from Paten 8.46e-01
20	28	1.4	1466	53	A07367	M.musculus VAC-alpha 8.46e-01
21	28	1.4	1560	78	HSVAC	Human mRNA for vascul 8.46e-01
22	28	1.4	1566	79	HUMBCI	Human blood coagulati 8.46e-01
23	28	1.4	1567	57	I33410	Sequence 2 from paten 8.46e-01
24	28	1.4	1573	44	RFAT2A711	R.fillosa mRNA for alp 8.46e-01
25	28	1.4	1575	53	AL3845	H.sapiens PP4 mRNA. 8.46e-01
26	28	1.4	1582	82	HUMPAP4	Human placenta antio 8.46e-01
27	28	1.4	1592	80	HUMENN	Human endonexin II MR 8.46e-01
28	28	1.4	1605	55	I07181	Sequence 12 from Paten 8.46e-01
29	29	1.4	2028	78	HSU56813	Human polycystin MRN 2.65e-01
30	29	1.4	2112	85	MMEBK	M.musculus mRNA for E 2.65e-01
31	29	1.4	2260	87	MMU43320	Mus musculus putative 2.65e-01
32	28	1.4	2290	43	PEAHSP	Plasmodium falciptarum 8.46e-01
33	28	1.4	3205	22	ECOUVRAA	E.coli uvra gene enco 8.46e-01
34	29	1.4	8595	39	DDU14576	Dictyostelium discoid 2.65e-01
35	28	1.4	10419	21	ECAE000479	Escherichia coli from 8.46e-01
36	28	1.4	176195	22	ECOUW89	E. coli chromosomal r 8.46e-01
37	27	1.3	105	56	I14734	Sequence 13 from paten 2.63e+00
38	27	1.3	1539	61	CRCPRPROM	C.roseus cpr gene, pr 2.63e+00
39	27	1.3	1738	60	ATU53866	Arabidopsis thaliana 2.63e+00
40	27	1.3	2147	63	NTPKTL7	N.tabacum PKTL7 mRNA 2.63e+00
41	27	1.3	3215	35	AGRNAG1	A.gambiae mRNA for im 2.63e+00
42	27	1.3	4203	85	MM5HT3RCB	M.musculus 5HT3 serot 2.63e+00
43	27	1.3	8631	44	PEPOLII	Plasmodium falciptarum 2.63e+00
44	27	1.3	27694	62	HASMT	Hansenula wingei mito 2.63e+00
45	27	1.3	62640	32	CEF26H11	*** SEQUENCING IN PRO 2.63e+00

108 (a)don(b)

26697

ALIGNMENTS

RESULT 1

LOCUS D50462 1776 bp mRNA ROD 19-DEC-1996
DEFINITION Mouse SDF5 mRNA, complete cds.
ACCESSION D50462
KEYWORDS SDF5.

REFERENCE 1 (bases 1 to 1776)
AUTHORS Shirozu, M.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-1995) to the DDBJ/EMBL/GenBank databases, Michio Shirozu, Kyoto University, Faculty of Medicine, Department of Medical Chemistry, Yoshida, Sakyo-ku, Kyoto, 606, Japan (E-mail: kondo@virus1.virus.kyoto-u.ac.jp, Tel: 81-75-753-4377, Fax: 81-75-753-4388)

REFERENCE 2 (bases 1 to 1776)

SOURCE

Mus musculus cell_line: ST-2 cDNA to mRNA, clone_lib: phage (lgt22a) library, ST-2 directional 1.

ORGANISM

Mus musculus
Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE

1 (bases 1 to 1776)

Shirozu, M.

Direct Submission

Submitted (01-MAY-1995) to the DDBJ/EMBL/GenBank databases, Michio Shirozu, Kyoto University, Faculty of Medicine, Department of Medical Chemistry, Yoshida, Sakyo-ku, Kyoto, 606, Japan (E-mail: kondo@virus1.virus.kyoto-u.ac.jp, Tel: 81-75-753-4377, Fax: 81-75-753-4388)

REFERENCE 2 (bases 1 to 1776)

[illegible]

Qy 579 CGAGACCATCCAGCCATGCCACTCGCTCTCGCTGACAGTGAAGGACCGCTGCGCCCGGT 638

Db 504 catgaacaagtctcggttttcagtgagccagagcctgcgctggagacacttccc 557

Qy 639 CATGTCGCGCTTCGGCTCCCTCGCTGCCGACATGCTTGATGAGTGCGACCGTTTCCC 692

RESULT 6

LOCUS I28278 215 bp DNA PAT 30-OCT-1996

DEFINITION Sequence 5 from patent US 5569830.

ACCESSION I28278

NID g1819054

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 215)

AUTHORS Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.

TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease

JOURNAL Patent: US 5569830-A 5 29-OCT-1996;

FEATURES

Location/Qualifiers

1..215

/organism="unknown"

BASE COUNT 15 a 8 c 25 g 26 t 141 others

ORIGIN

Query Match 1.9%; Score 39; DB 57; Length 215;

Best Local Similarity 14.7%; Pred. No. 7.33e-07;

Matches 24; Conservative 72; Mismatches 64; Indels 3; Gaps 3;

Db 43 vcdtdttrvnnndshnkysanynngnnvgaakthyyth-tnvsdgadsktvdtnas 101

Cp 547 GCGAGCACAGGAACCTCTGGTGTCCGGTGGCAGCTGCTCATGACACGCGGATCCAG 488

Db 102 gtssnggtgdnrs-gadysgasktamtrnrtgkannavdrnmngdasvsgdntkxh 160

Cp 487 CGCCGCGCTTCCAGACCTCCTCATGCTGTCTGGCCGACGAGTGTGGCAGCGCA 428

Db 161 aknsadgkvsgknnrdnnrnyrtgtksvsnncggnkrdvss 203

Cp 427 TGTTCGTGATTCGATCGCGTG-GCACAGCTGACAGTTGGCCG 386

RESULT 7

LOCUS RATER22H 4540 bp mRNA ROD 16-JUL-1993

DEFINITION Rattus norvegicus Drosophila polarity gene (frizzled) homologue mRNA, complete cds.

ACCESSION L02529

NID g310116

KEYWORDS frizzled gene; homologue; polarity gene.

SOURCE Rattus norvegicus (strain Sprague-Dawley) osteosarcoma cDNA to mRNA.

ORGANISM Rattus norvegicus

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Rattus

REFERENCE 1 (bases 1 to 4540)

AUTHORS Chan,S.D.H., Karpf,D.B., Fowlkes,M.E., Hooks,M., Bradley,M.S., Vuong,V., Bambino,T., Liu,M.Y.C., Arnaud,C.D., Strewler,G.J. and Nissenon,R.A.

TITLE Two homologs of the Drosophila polarity gene frizzled (fz) are widely expressed in mammalian tissues

JOURNAL J. Biol. Chem. 267, 25202-25207 (1992)

MEDLINE 93094228

FEATURES

Location/Qualifiers

1..4540

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/cell_line="UMRI06-01"

/cell_type="osteoblast"

/tissue_type="osteosarcoma"

CDS

766..2691

/note="Drosophila polarity gene (frizzled) homologue"

/codon_start=1

/db_xref="pid:g310117"

/translation="MAEAVPSGESRAAGRPSELCALVALPGRREEVGHQDTAGHRRPR

AHSRCWARGLLLLWLEAPLLGVRAQAGVSGQORPPPPPOQGOQVNGERG

ISIPDHGYCQIPISIPCTDIANNQTIMPNLLGHTNQEDAGLEHVQYPLVKVQCSAEL

KFELSMYAPVTVLEQALPPCRSLCERAGCCALMNKFGQMPDILKCKFPVHGAG

ELCVQNTSDKGTPTPSLLPEFWTSNPQGGGGRGYPGGAGVGRGKFCSPRALRV

PSYLNHFLEKDCGAPCEPTKVYGLMYFGPEELRFSRTWIGIWSVLCCLASTFTVLT

YLVDNRRFSYPERPIIFLSCGYTAVAVIAGLELDRVVCNDKFAEDGARTVAOGTK

KEGCTILENMLYFFSMASSIMVILSLTWFLAAGMKMGHEAIEANSQYFHLAAWVPA

IKTITILALQYDGDVLSGCVFGLNNVDALRGFVLAPLEVILFISTFELLAGEFVSLF

RIKTMKHGTQTEKLEKLMVRIGFVSLITVATVACIFYEQAFRDQWERSWVAQ

SCKSYAIPQPHLQGGGVPHPMPSPDFTVMKILMTLITVITSGTGFVWSGRTLSNQ

RKFYTRLNKSQGETTV"

BASE COUNT 998 a 1255 c 1220 g 1067 t

ORIGIN

Query Match 1.9%; Score 38; DB 90; Length 4540;

Best Local Similarity 61.6%; Pred. No. 2.85e-06;

Matches 181; Conservative 0; Mismatches 107; Indels 6; Gaps 5;

Db 1116 gctgtgacggagacatcggtacaaatcagaccatcatcgcacacacgtgtgggcccacagaa 1175

Qy 399 GCTGTGCCACGGCATGAATACCAGAACATCGCGCTGCCCAACCTGCTGGGCCACGAGAC 458

Db 1176 tcaggagagcgccgctgaggtgcaccagtctcaccgttctcaccgttctgaggtgcagtgctc 1235

Qy 459 CATGAGGAGGTGCTGGAGGAGCGCCGGCTTGGATGCCCTGCTCATGAGCAGTGCCTCA 518

Db 1236 agcggagctcaagtcttctcctgtgctccatgacgcgcctgtgtgca-cggt-ac-tgga 1292

Qy 519 CCGGACACCAAGAAGTCTCTGCTGCTGCTTTCGCCCGCTGCTGCTCATGACCTAGA 578

Db 1293 gcaggcgctgctcctcctgcgcctcctcctgctgctgagcgcgccagggctgc-gaggcac--t 1349

Qy 579 CGAGACCATCCAGCATCGCACCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 638

Db 1350 catgaacaagtctcggttccagtgccagacacacacacacacacacacacacacacacacac 1403

Qy 639 CATGTCGCGCTTCGGCTTCCCTGCGCCGACATGCTTGAAGTGCGACCGTTTCCC 692

RESULT 8

LOCUS I28278 215 bp DNA PAT 30-OCT-1996

DEFINITION Sequence 5 from patent US 5569830.

ACCESSION I28278

NID g1819054

KEYWORDS .

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 215)

AUTHORS Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.

TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease

JOURNAL Patent: US 5569830-A 5 29-OCT-1996;

FEATURES

Location/Qualifiers

1..215

/organism="unknown"

BASE COUNT 15 a 8 c 25 g 26 t 141 others

ORIGIN

Query Match 1.5%; Score 31; DB 57; Length 215;

Best Local Similarity 11.0%; Pred. No. 2.40e-02;

Matches 22; Conservative 82; Mismatches 95; Indels 1; Gaps 1;

Db 4 vtmsssvsvrtasondkakkdnttsswttdcncrtwgvcddttrvnnndshnkys 63

Qy 382 ATCCCGGCCAACCTGCGAGCTGTGCCAGGCGATCGAATACCAGAACATGCGGCTGCCCAAC 441

[illegible]

Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
Charophytica/Embryophyta group; Embryophyta; Magnoliophyta;
Magnoliopsida; Caprales; Brassicaceae; Arabidopsis.
1 (bases 1 to 4297)
King.K.
Unpublished
2 (bases 1 to 4297)
King.K.
Direct Submission
Submitted (22-JAN-1994) K. King, Dept of Genetics, Universitaet
Tuebingen, Auf der Morgenstelle 28, 72076 Tuebingen, FRG

FEATURES
source location/Qualifiers
1..4297
/organism="Arabidopsis thaliana"
/strain="Ostwestfalen"
/tissue_type="leaf"
/clone="Athsc70g"
<1..171
/gene="Hsc70-1"
/number=1
join(<1..171,496..2238)
/gene="Hsc70-1"
/codon_start=1
/product="heat shock cognate 70-1"
/db_xref="PID:g450880"
/db_xref="SWISS-PROT:P22953"
/translation="GTYSVCGVQWHDREVIIANDOGNRTTPSVVAFDTSERLIGDAA
KNQAVSPVNTVFDAKRLGRFSDSSVQDMKLWPFKIOAGPADKPMIVYVKGGEK
EFAAEISNVLLIKMEIEAAYLGVTIKNAVTVPAVFENDSQRAQTMDAGVIAGLVN
RIINEPTAAIAGLDKKATRVSEKNVLIFDLGGTFDVSLLTIEGIFEVKATAGDT
HLGEGEDNRNVHFVOEFRKRKKDDINPRALRRLRSCERAKRTLSTSAQTTEI
DSLYIEDGFYSTITRAREFEELNDLFRCMPEVEKCLRDKMDKSTVHVYLVGSGTR
IPKVQLLQDFNGKELCKINSPIEAUVYAAGVQALLSGEGNEKVKVDLLLDVPLPS
LGLETAGGVMTTILPNTTPTTKEQVFSTYSDNOPGLVLIQVVEGERARTKDNLLGK
FESLGPAPARGPVQITVCFDIDANGILNVSAEDKTGQKNKYTIINDKGLSKDELE
KMVQEAEKYSDEEHKKVEAKNALENYNMRNTIQDEKIGEKLPAAOKKKIEDISI
EQAIQWLQEGNLGADEDEFDMKKELESICNPITIAKMYQQAGEAGPGGASGMDDADDP
ASSGAGPKIEVD"
172..495
/gene="Hsc70-1"
/number=1
496..>2238
/gene="Hsc70-1"
/number=2
<3695..3908
/gene="Hsc70-2"
/note="see also M23106"
/number=1
join(3695..3908,4217.>4297)
/gene="Hsc70-2"
/note="see also M23106"
/codon_start=1
/product="heat shock cognate 70-2"
/db_xref="PID:g450881"
/db_xref="SWISS-PROT:P22954"
/translation="MAGKGKGAIGIDLGTYSVCGVQWHDREVIIANDOGNRTTPSY
VATDTSERLIGDAAKNQVAMNPVTFDAKRLIGRRFSASVQSQRQLWFPTII"
3909..4215
/gene="Hsc70-2"
/note="see also M23106"
/number=1
4217..>4297
/gene="Hsc70-2"
/note="see also M23106"
/number=2

BASE COUNT 1172 a 840 c 906 g 1379 t

ORIGIN

Query Match 1.5%; Score 30; DB 59; Length 4297;
Best Local Similarity 65.0%; Pred.No. 8.07e-02;
Matches 65; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Db 2982 gcctatttttttttttggaacacctgatcctttgggttatagcttgatgttgtaat 3041

Cp 2021 GCTTTT|||||TTTTTT|||||TTTTTTAGGTAAACAGGATGTAAAGCTTTTATACAAGAATAATG 1962
Db 3042 ttatctgggttttaccttgacgttttatcgtttattt 3081
Cp 1961 TTTATCGAATAATTACAGTGTTGGTTAAAGCAATATTT 1922

RESULT 12
LOCUS CELK02A2 38262 bp DNA INV 13-JUL-1995
DEFINITION Caenorhabditis elegans cosmid K02A2.
ACCESSION U23171
NID 9726381
KEYWORDS
SOURCE
ORGANISM
Caenorhabditis elegans strain=Bristol N2.

Eukaryotes; mitochondria eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabdittia; Rhabdittida; Rhabdittina; Rhabditoidea;
Rhabdittidae; Caenorhabditis.

REFERENCE 1 (bases 1 to 38262)
AUTHORS Wilson,R., Alnough,R., Anderson,K., Baynes,C., Berks,M., Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J., Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A., Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M., Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N., Letrait,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B., O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A., Saunders,D., Showkeen,R., Smaldon,N., Smith,A., Sonhammer,E., Staden,K., Sulston,J., Thiermy-Mieg,J., Thomas,K., Vaudin,M., Vaughan,K., Waterston,R., Watson,A., Weinstock,L., Wilkinson-Sproat,J. and Wohlman,P.
TITLE 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans

JOURNAL Nature 368 (6466), 32-38 (1994)
MEDLINE 94150718
REFERENCE 2 (bases 1 to 38262)
AUTHORS Hallsworth,K.
TITLE The sequence of C. elegans cosmid K02A2
JOURNAL Unpublished (1994)
REFERENCE 3 (bases 1 to 38262)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (13-JUL-1995)

SUBMITTED BY:
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RO, England
e-mail: rwenematode.wustl.edu and jess@sanger.ac.uk

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is C07D10, 200 bp overlap, 3' cosmid is T19C12.
Actual start of this cosmid is at base position 197 of CELK02A2;
actual end is at bp 38262 of CELK02A2

NOTES:

Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green and L. Hillier, ms in preparation).

FEATURES source Location/Qualifiers
1..38262
/organism="Caenorhabditis elegans"
/map="I1"
/strain="Bristol N2"
join(2467..2528,2604..2700,2748..2885,3030..3083,3384..3656,3703..3804,4217..4404)
/gene="K02A2"

CDS

```

/note="similar to C. elegans protein E04F6.1"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g726382"
/translation="MYGVCFHCKNLSMTEKIKNLSIPNIDSFVAVLPGRPGREFGPDV
CFSYHFFGLVCLVGLGTSNIIIFRQALRGVRSRIFLTMISLTYIPSSITIVYL
NYFLRKNLNLTSOILPSTSDPEKESLVLYIEHPKYDLSTIEFVGFHNIAFQF
TLATLLLVIGAVAIAPISGFLSRVHLINDNRGSLKTEQSKTILAYGLACQTFPLV
ICVYPASCIIFSQMSTVELLNEHLGLICFPSPDFISFYFIVPYRQALLGLFK
CKRTPLENIIVIRHLSKRNPISVDSN"
complement(join(4846..4918,4971..5077))
/genes="K02A2.5"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g726384"
/translation="MLSDDFSLFSYDDKADAHDAADAAFRKFEALCSARSQOPSTSK
PSLIDALCTSAKKE"
complement(join(19741..19827,19873..20071,20408..20658,
20735..21024,21067..21346,21423..21621,21674..21793,
21924..22110,22353..22900,22950..23072,23118..23551,
23600..23743,23809..23993,24055..24070))
/genes="K02A2.3"
/note="similar to bumetanide-sensitive Na-K-Cl
cotransporter (PIR:A53491)"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g726386"
/translation="MGKGGRMETGEGDIPADYKGNRKFTHSLGHLYLKEDEGIGT
QASFISGTYTPGPKERATSEHVKANGLVGLVPTIOHLGVTMEIRLFWVNGSV
ANTMALLAICCLSTLLTSLSNAVATNGVVGSGAVFIISRLGAEFGAGILFYLA
NTVASMIVGGVEVILMYLPEMAIGGADALHDTMFGLSNLYNRLDITVELLIQAL
IVAMGVFQQLAPSLCMVILAIACIGGIEKQITMEGMYKALIDNHLLOSSIVTH
PIHKNOTSFNETFCNCLSKLYLSEFCANVNDEASAEADVCTHTVTKKMTQC
LAFPGNKTNDNMWPEKSESVVPGVGRKETAEOVODESTFPLMAIYFPAVYG
ITFGNMGDLRDOPRSIPVGTIAATLTTSALYILAILFGSITSVLRDKRGSIG
NMYVVALSWPAPVTVTGAFSLTGAALQCLCSAPRLLQSIKANDVIILAPPARVT
KNEPFLGLWPIVIAECGILLGAVDIAEVLDFFLMTLSLGLAECFFIMFASVP
LACIATAVAYIKVYEWKAKKMGDGRGALATAQYLSLKVEDKDPHPKNRPQV
LITLLTSQSKEMIDRRVAMNLGKAGKAGRLATACAFKLGSDVKDKNARDVKT
TLVKDMSVRLRGFAKTFMEYNNHQINGTISGLVQSIGIGLRPNTILLNWPNEKNPDE
LVLPAAEIIHGAANDCLVTGIDTPEYSERLTGFIWMIVODGGILMLIAYILR
QHKVKGCTLRIFAVSEQSTSKEMKAGLQYIYMLRIDAEILFVLDLMEVSDDEV
EKAAVERKQRESEMRKSGYLDNFGMEDNGKPRQVMRHSDSARSFSPQGAHTS
INDLETSTFESLEDFYRGTPNEDELEGAMKLNHKMNTSVLRNVRVRENSPDSQL
ILLNPSPPRNLAFNNSYMTYLDLTEDLPRLVFIGGSGREVITIDS"
join(26353..26567,26616..26712,26757..26894,26725..29289,
29341..29441)
/genes="K02A2.1"
/note="similar to C. elegans protein E04F6.1"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g726385"
/translation="MSDRSNYSNRFVSYNAYHFFLVLVAGLTFQCILLYLIRTKSPAS
LDQLKIFLNTAFVLIVICIGFTOHRSLPNSSTPFAVLANGPCRVFGTIVCGGHV
FLGVTFVALISNTVYIFLLRRDITKHLIYIILSYVPGIITVAPLLILSFN
LVSVLFDVTYFTELKLAQEHGYVALVDIYKGRICTDISSAVTLLRPMISDR
NGKLPRLEALNALKSPVCKQKLGAFGFCIGGLCSLDCARYRDGTRAVTSFHTG
LTPIEGLPDLDDNSIOVHGDAKHVSKVTVDAPHEEMRARNSDVFVISHGKAMHS
FIDPSAGIAPSGVGDANAEKRSKATUEFLKEAFA"
complement(join(29971..30119,30168..30403,30450..30655))
/genes="K02A2.4"
/note="coded for by C. elegans CDNA CEE5P19F; similar to
C. elegans protein E04F6.1"
/codon_start=1
/db_xref="PID:g726383"
/translation="WADDSNRFVLFDVYVYYSFIIGVFAQSVLLYLKKNKSPRIYS
FCYFINTWIVQFALLMTFFTSQCLPNTSTYTVLPGRGPFYGTAGFAGYHSLA
YSMAVALGFANTVLFYRLPIRFHGNKLFDDDCMLHSSNCDCTSFGLHGFSGS
TYLHLPRISDVFEENRTISLVLRISNQHFSLRL"
complement(join(33175..33695,33825..35082,35609..37636))
/genes="K02A2.6"
/note="similar to silkwork transposon mag (PIR:S08405) and
POL polyprotein (reverse transcriptase)"

```

```

/codon_start=1
/evidence=not_experimental
/db_xref="PID:g726387"
/translation="MNDLRKSQISQDIYNSFDDTVALTKIFGTSKSLMRRQOCQIQI
TRILLCLCESTFHRYSQDESDREISFDFDVALTKIFGTSKSLMRRQOCQIQI
CRANGSLQDYLDYNTSIDAVLSDKSSMTSDWSIFLFLRGLNSPDGEKAKLYLMQY
VPASEKKNKELKLSLDVHDEWKFLOHQOQSKIVSVKPSKSSQOVDNVKVDNRSKKK
KPIPRKPESSQSKKGEIPCYCNKKGHYATNCRSPKTNQNGKNGKSKGCSYV
HVDGLVKTEHQAKHRMSVEVCGKDAFQDGTGSMITLISVKWCKLSPPLKVPKPHR
ISCANTPMVAKRGCLVKFKLGIEYEVYVDRDNLGTGKKAETEEENAVPVFKRAPVPI
VQVSTSETEASLEVMKNDPFEKDGGLGCKEKAETEEENAVPVFKRAPVPI
GSLEAVETELNRLQEMGVLPITYAKWAAPIVIKKGTGKIRVCADKCSGLNAALK
DEHPLPTSEDIFSRLKGVYSQIDLKDAYLOVELEDEEAKGLAVINTHRCIFKLRMT
FGKLPAFPASFKIMDKMVSGLTGVAVWDDIILISASIEHEKILRELPERFKEYGR
VSAEKAFAKQKVTFGLGVDEHRRPDSKRTAISRMAKPTDQKQASFLGAADWLRS
MMQDHQONADDVVIAYIDDDDDDDSIQKLNPTVETIDIRFSQKDHESVSVKLYR
NDSWKPKSTEIEKHWRDLRLKHLHGCLLDLDDRVTPKSLQKIVLKLQHEHGPVIV
OMKQKARSFVFWGLDSDIENMVHCHNCQENSKMPRVVPLNPNVPEAPWKRIHIDF
AGPLNCYLLVVVDATKTYAEVKLTRISAVTTIDLEIFSHGYPETIISNGTQL
TSHLFAQMCOSGIEHKTSAVYIPRSGAAREVDLLKGIKIKEGSGVNOQLLNF
LISYRNTPHSALNGSTPAECHFGKIRITMSLLMPTDRVLKPKLTQYQNMKHHYEL
RNGARAKAFQVQKVVYVHGHKQWKHGVIRKFGVGLYEVQVGDGMQKSHVQIIR
TRYGDIYSSRVNPLSSDPLGFVGGSGDEMNRDDPVVTSNGSSVDVNRGSRITSEL
KKESNAVRGQPCGSCSPNNDSAPGYASRSHPTTDPSSHVSRSRIRREVLDRYSS
VEHPSTSTGTPRGSTSTQLQGASTRNGSRITAGSRNPSCQGNRISSIRGEGVTARRR
VRTWR"
BASE COUNT 12820 a 6635 c 6904 g 11903 t
ORIGIN
Query Match 1.5%; Score 31; DB 37; Length 38262;
Best Local Similarity 73.1%; Pred. No. 2.40e-02;
Matches 49; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
Db 20220 aaagcgaaacttcacgaacttatagaataatcaataaataaactgagaattt 20279
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1346 AAAGGAAAACCCACCACCATCTTGTAGAAATATTCAAACTATAAATCATGAATATT 1405
Db 20280 ttctcgaa 20286
||| |||
QY 1406 TTATGAA 1412
RESULT 13
ID OFU89259 standard; DNA; INV; 354 BP.
AC U89259;
NC q1881675
DT 13-MAR-1997 (Rel. 51, Created)
DE 13-MAR-1997 (Rel. 51, Last updated, Version 1)
DE Oxytricha fallax 57kd zinc finger/protein chimera gene, partial
DE cds.
KW Oxytricha fallax
OS Eukaryotae; mitochondrial eukaryotes; Alveolata; Ciliophora;
OC hypotrichs; Sticotrachida; Oxytrichidae; Oxytricha.
RN [1]
RP 1-354
RX MEDLINE; 94134747.
RA Doak T.G., Doerder F.P., Jahn C.L., Herrick G.;
RT "A proposed superfamily of transposase genes: transposon-like
RL elements in ciliated protozoa and a common 'p35E' motif";
RL Proc. Natl. Acad. Sci. U.S.A. 91:942-946(1994).
RN [2]
RP 1-354
RA Witherspoon D.J., Doak T.G., Williams K., Seger J., Herrick G.;
RT "Selection on the protein-coding genes of the TBEL family of
RT transposable elements in the ciliates Oxytricha fallax and O.
RL trifallax";
RL Unpublished.
RN [3]
RP 1-354
RA Doak T.G., Williams K., Witherspoon D.J., Herrick G.;
RT ;
RL Submitted (11-FEB-1997) to the EMBL/GenBank/DBJ databases.

```



```
/note="lipocortin V, endonexin II, PAP-I, PP4, VAC-alpha"
/codon_start=1
/product="annexin v"
/db_xref="PID:9430966"
/translation="MAQVLRGTVDFPGFDERADAEITLRKAMKGLGTDEESILTLTTS
RSNAQRQISAAFTLFGRLDLDLSELTGKFEKLI VALMKPRLYDAYELKHALKG
AGTNKQVTEITASTPEELRAIKOVYEEYSGSLEDVVGDTSGYORMLVLLQAN
RPPDAGIDEAOVEQDAQALFOAGELKMGCTDEKFTIFGTRSVSHLRKYVDKYMVISG
FOIETIDRETSGNLEQLLLAVVKSIRSIPAYLAETLIYAMKGAGTDDHTLIRVWVSR
SEIDLFNIRKFRKNFATSLYSMIRGDTSGDYKAKLLLLLCGEDD"
intron
<1..30
/evidence=experimental
/gene="ANX5"
/note="full size 0.8 kb; phase 0 intron"
/exon
/number=12
31..565
/evidence=experimental
/gene="ANX5"
/number=13
91..565
/evidence=experimental
/gene="ANX5"
538..543
polyA_signal
/gene="ANX5"
BASE COUNT 163 a 107 c 104 g 191 t
ORIGIN
Query Match 1.4%; Score 28; DB 73; Length 565;
Best Local Similarity 80.4%; Pred. No. 8.46e-01;
Matches 37; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Db 345 atatattttttaagagcctagagcgttttagccttttttaaaa 390
Oy 1794 ATATAAATTTTATAATTGACTTGAGTACTTTAAGCCTTGTTTAAAA 1839
```

Search completed: Tue Dec 9 08:36:17 1997
Job time : 2317 secs.

W P E R E H (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 09:45:55 1997: MasPar time 699.37 Seconds
Tabular output not generated. 1295.567 Million cell updates/sec

Title: >848439-3-trans
Description: (1-825) from translate.seq
Perfect Score: 688
N.A. Sequence: 1 WSNCGNMGNGNYTNTTYT.....SNATNMGNAARYTNCARTGY 825
Comp: SWNCGNKCNCNRANAARRA.....WNTANKNTTYRANGTYACR

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0: Query 0

Searched: 362067 seqs, 549138275 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embli-new3
1: BCT 2: FUN 3: GEN1 4: GEN2 5: HTG1 6: HTG2 7: HUM 8: INV
9: ROD 10: NAM 11: VRT 12: PLN 13: PRO 14: ROD 15: SYN 16: UNC
17: VIR

Database: genbank99
18: BCT1 19: BCT2 20: BCT3 21: BCT4 22: BCT5 23: BCT6 24: BCT7
25: BCT8 26: BCT9 27: BCT10 28: BCT11 29: GEN1 30: GEN2
31: GEN3 32: HTG1 33: HTG2 34: HTG3 35: INV1 36: INV2 37: INV3
38: INV4 39: INV5 40: INV6 41: INV7 42: INV8 43: INV9 44: INV10
45: INV11 46: NAM1 47: NAM2 48: NAM3 49: VRT1 50: VRT2 51: VRT3
52: VRT4 53: PAT1 54: PAT2 55: PAT3 56: PAT4 57: PAT5 58: PHG
59: PLN1 60: PLN2 61: PLN3 62: PLN4 63: PLN5 64: PLN6 65: PLN7
66: PLN8 67: PLN9 68: PLN10 69: PLN11 70: PRI1 71: PRI2
72: PRI3 73: PRI4 74: PRI5 75: PRI6 76: PRI7 77: PRI8 78: PRI9
79: PRI10 80: PRI11 81: PRI12 82: PRI13 83: PRI14 84: PRI15
85: ROD1 86: ROD2 87: ROD3 88: ROD4 89: ROD5 90: ROD6 91: ROD7
92: ROD8 93: SYN 94: SYN 95: UNA 96: VRL1 97: VRL2 98: VRL3
99: VRL4 100: VRL5 101: VRL6 102: VRL7 103: VRL8 104: VRL9
105: VRL10
Database: genbank-new3
106: BCT 107: GEN1 108: GEN2 109: HTG1 110: HTG2 111: INV
112: NAM 113: VRT 114: PHG 115: PLN 116: PRI1 117: PRI2
118: ROD 119: SYN 120: UNA 121: VRL
Database: u-embli50_99
122: part1

Database: genbank-new3
106: BCT 107: GEN1 108: GEN2 109: HTG1 110: HTG2 111: INV
112: NAM 113: VRT 114: PHG 115: PLN 116: PRI1 117: PRI2
118: ROD 119: SYN 120: UNA 121: VRL
Database: u-embli50_99
122: part1

Statistics: Mean 14.925; Variance 11.240; scale 1.328

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

3

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	682	99.1	1776	85	D50462	Mouse SDF5 mRNA, comp	0.00e+00
2	89	12.9	2421	87	MMU43321	Mus musculus putative	5.75e+24
3	75	10.9	2334	77	HSU43318	Human putative trans	2.82e+17
4	64	9.3	1770	38	CEU43316	Caenorhabditis elegans	3.27e+12
5	63	9.2	2831	87	MMU43317	Mus musculus putative	9.21e+12
6	61	8.9	2085	40	DMU65589	Drosophila melanogaster	7.21e+11
7	59	8.6	2260	87	MMU43320	Mus musculus putative	5.54e+10
8	56	8.1	1913	38	CEU63557	Caenorhabditis elegans	1.13e+08
9	56	8.1	4540	90	RATPR22H	Rattus norvegicus Dro	1.13e+08
10	54	7.8	1923	80	HUMPR12	Human frizzled gene p	8.23e+08
11	52	7.6	1912	90	RATPR2H	Rattus norvegicus Dro	5.83e+07
12	52	7.6	34576	8	CET23D8	Caenorhabditis elegans	5.83e+07
13	52	7.6	34576	38	CET23D8	Caenorhabditis elegans	5.83e+07
14	52	7.6	34576	111	CET23D8	Caenorhabditis elegans	5.83e+07
15	49	7.1	215	57	I28278	Sequence 5 from paten	1.05e+05
16	47	6.8	215	57	I28278	Sequence 5 from paten	6.91e+05
17	46	6.7	2624	87	MMU43205	Mus musculus frizzled	1.76e+04
18	44	6.4	354	8	OFU89259	Oxytricha fallax 57kd	1.10e+03
19	44	6.4	354	111	OFU89259	Oxytricha fallax 57kd	1.10e+03
20	40	5.8	354	111	OFU89259	Oxytricha fallax 57kd	3.89e+02
21	40	5.8	354	8	OFU89259	Oxytricha fallax 57kd	3.89e+02
22	39	5.7	74911	33	HS209H1	Human DNA sequence **	9.24e+02
23	37	5.4	1509	23	INSREP	Insertion sequence IS	5.03e+01
24	37	5.4	1513	24	MMU01217	Mycoplasma hyorhinis	5.03e+01
25	37	5.4	1518	23	INSGDLA	Insertion sequence IS	5.03e+01
26	37	5.4	2187	43	FFAHSPT70H	Plasmidium falciptarum	5.03e+01
27	36	5.2	2492	87	MMU43319	Mus musculus putative	1.15e+00
28	36	5.2	140571	34	HS46H23	Human DNA sequence **	1.15e+00
29	35	5.1	1291	11	XLU78598	Xenopus laevis Frzb p	2.60e+00
30	35	5.1	1291	113	XLU78598	Xenopus laevis Frzb p	2.60e+00
31	35	5.1	1875	52	XLU68059	Xenopus laevis frzlp	2.60e+00
32	35	5.1	4284	24	MHVLPABC	M.hyorhinis vlpa, vlp	2.60e+00
33	35	5.1	33200	36	CEF34B3	Caenorhabditis elegans	2.60e+00
34	35	5.1	47745	33	CET26E3	*** SEQUENCING IN PRO	2.60e+00
35	34	4.9	413	44	PFU17812	Periplaneta fuliginos	5.79e+00
36	34	4.9	1054	39	DDU67923	Dictyostelium discoid	5.79e+00
37	34	4.9	3618	38	CEZYG11	Caenorhabditis elegans	5.79e+00
38	34	4.9	4413	69	YSPCRM1N1	Yeast crml-N1 gene fo	5.79e+00
39	34	4.9	7698	35	CCU60529	Ceratitidis capitata yo	5.79e+00
40	34	4.9	13815	88	MUSDYSA	Mouse dystrophin mRNA	5.79e+00
41	34	4.9	26321	35	CEC08B11	Caenorhabditis elegans	5.79e+00
42	34	4.9	34731	78	HSU75931	Human Xp22 cosmid U25	5.79e+00
43	34	4.9	39569	37	CELE01H11	Caenorhabditis elegans	5.79e+00
44	34	4.9	39874	38	CEZK287	Caenorhabditis elegans	5.79e+00
45	34	4.9	119704	63	OSCHLPLXX	O.sinensis chloropias	5.79e+00

ALIGNMENTS

RESULT	1	D50462	1776 bp	mRNA	ROD	19-DEC-1996
LOCUS		Mouse SDF5 mRNA, complete cds.				
DEFINITION		D50462				
ACCESSION		91747301				
NID		SDF5.				
KEYWORDS		Mus musculus cell_line:ST-2 cDNA to mRNA, clone_lib:phage (lgt22a)				
SOURCE		library, ST-2 directional 1.				
ORGANISM		Mus musculus				
		Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;				
		Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
		Mus.				
REFERENCE	1	(bases 1 to 1776)				
AUTHORS		Shirozu,M.				
TITLE		Direct Submission				
JOURNAL		Submitted (01-MAY-1995) to the DDBJ/EMBL/GenBank databases. Michio				
		Shirozu, Kyoto University, Faculty of Medicine, Department of				
		Medical Chemistry; Yoshida, Sakyo-ku, Kyoto, Kyoto 606, Japan				
		(E-mail: kondo@virus1.virus.kyoto-u.ac.jp, Tel:81-75-753-4377,				
		Fax:81-75-753-4388)				
REFERENCE	2	(bases 1 to 1776)				

BASE COUNT	356 a	803 c	736 g	439 t
ORIGIN				
Query Match	10.9%	Score 75;	DB 77;	Length 2334;
Best Local Similarity	39.7%;	Pred. No. 2.82e-17;		
Matches	116;	Conservative	44;	Mismatches 132; Indels 0; Gaps 0;
Db	439	ttgtgcgcggcatcggtctacaacactgacgcacatgcccaaccagttcaaccacgacgc	498	
QY	86	TNTGYCAYGNATNGARTAYCARAAATGMGNTNCCNAAAYTNTYNGGNCAYGARACNA	145	
Db	499	aggacagggcggcctgaggtgcacacgttctggcgcgtggtggagatccaatgctgc	558	
QY	146	TGAARGARTNTYNGARCARGCNGGNCNTGGATNCCNTYNGTNATGAARCAFTGYCAYC	205	
Db	559	cggacgtgcgctcttctctcatgcatctacacgcccctctgtctgcgcactaccaca	618	
QY	206	CNGAYACNAAARAARTTYTNTGYWSNYTNTYGCNCCNGTNGYTYNGAYGAYTYNGAYG	265	
Db	619	agcgcgtgcgcctcgcgcctgcggtgtgcgagcgcgccaaaggcgcgcgtcgtcgccgtga	678	
QY	266	ARACNATNARCCTNGCYAWSNYTNTGYTCNARGTNAARGAYMGNTGYGCNCCNGTNA	325	
Db	679	tcgcgcagtcacgcttcgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcctccc	730	
QY	326	TGWSNCGNTTYGNTTYCCNTGGCCNGAYATGYTNGARTGYGAYMGNTTYCC	377	
RESULT	4	CEU43316 1770 bp DNA INV 24-FEB-1996		
LOCUS		Caenorhabditis elegans putative transmembrane receptor (frizzled 1)		
DEFINITION		gene, complete cds.		
ACCESSION		U43316		
NID		g1151249		
KEYWORDS		Caenorhabditis elegans.		
SOURCE		Caenorhabditis elegans		
ORGANISM		Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea; Rhabditidae; Caenorhabditis.		
REFERENCE		1 (bases 1 to 1770)		
AUTHORS		Wang,Y., Macke,J.P., Abella,B.S., Andreasson,K., Worley,P., Gilbert,D.J., Copeland,N.G., Jenkins,N.A. and Nathans,J.		
TITLE		A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled		
JOURNAL		J. Biol. Chem. 271 (8), 4468-4476 (1996)		
MEDLINE		96224032		
REFERENCE		2 (bases 1 to 1770)		
AUTHORS		Abella,B., Wang,Y., Macke,J.P. and Nathans,J.		
TITLE		Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205		
JOURNAL		Location/Qualifiers		
FEATURES		source		
		1..1770		
		/organism="Caenorhabditis elegans"		
		/chromosome="I"		
		/map="on YACs Y95B8 and Y34D9"		
		57..1634		
		/gene="frizzled 1"		
		/note="fz 1; putative transmembrane receptor"		
		/codon_start=1		
		/product="transmembrane receptor"		
		/db_xref="PID:g1151250"		
		/translation="MGPRGYLVTWLLLLLVIGVGQRQKQKVDHEMCMNDLPYNTSFPNLVDESWKDAESILTKPLLSVYCSQKFLCYVIFPMCKLANPFGCPRLPLC LSVQEKCLPVLESGFKWPDTRCDKFPLENNREKCMKMGNEQAIQDERAKFAAKE SEDGNDVEDIQREVDLNGKCPQDEVLRSSKCVPLCSNPQKQVQDTRDSATRLL LFLSLSVLLILSVFIYGLSRLEMLHSITETAFPCISCATSVIIVISFQKQF QICSDTYTHLLLFVVGUSHVPCSVASLIITYTCSRLLMLLICVSNWKAFTSRHIL DDSRTVRIMLIGLIPAPLMALAKAAVAANPLTGLCFIGAASPGTDIFNFCRLIL FLTISLALLSSACCLLGGDSVNGFAGVIAAVYPIAGLYMLSFVNDATQPFLLDLR SFNAVSATKFSFDLLLLSFIMCAFLIYLLFKLTRSSSKSVKEGYQPAVPKLPQPAIPG		
CDS				

misc_feature		/chromosome="I"		Two homologs of the Drosophila polarity gene frizzled (fz) are widely expressed in mammalian tissues	
CDS		/note="S1 trans-splice leader"		J. Biol. Chem. 267, 25202-25207 (1992)	
		/gene="lin-17"		Location/Qualifiers	
		/function="regulates certain asymmetric cell divisions"		1..4540	
		/note="similar to Drosophila Frizzled"		/organism="Rattus norvegicus"	
		/codon_start=1		/strain="Sprague-Dawley"	
		/product="transmembrane receptor LIN-17"		/cell_line="UMR106-01"	
		/db_xref="PID:gl589780"		/cell_type="osteoblast"	
		/translation="MMHSLGIITFLPIATGSIPOAVKVKCIPIDIELCKDLPPYNT YPNTILHNDQHTLOTHTEHFEPKMKTKCHPHIFPICSVPAPMCPGIMQVATSCAS VCEQVAKCFSILEFEGWPEPLNCAQFPDPELCLMCAPTDEIDTGGFSAPRLPTKGS SSSSKPTCSDLVDVDPDHPDKSHCAFCACOSNMFSTDNKRMVRSINFWAANAGV ALFTFDKTRFDFRPERCVFYLSLCIFISSLPYTFPLFDIAPIRSCHALNGRSGY LGIGTFDNSYCLASLLNFFVFAALWMLFSFTLYSGGRKWPGEIACSVGHV VAWGLSSLIATIVLNFKNVDASELALCVGNLSIALWFIWIVPRICVIGPCFIV GFASWCRERISFRTRGDTSTKLEKLVKMGFFCALFILPNVIELVQCCKFPMILITQW TMTIDCKQSGACHRRPFPQAEIYTAVISSLATGFSCLMWLSAKTAVHAKNIFEC GMSCAQVKNKPIETSRPLLEPTTAPPPPPYVQMOMTNPQNSWRPSKV"		/tissue_type="osteosarcoma"	
		/db_xref="PID:g310117"		/note="Drosophila polarity gene (frizzled) homologue"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/codon_start=1	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKFYRLRNSKQGETTV"		/db_xref="PID:g310117"	
		/translacion="MAEAVPSESRAAGRPSSLECAVALPGRREVEGHODTAGRRRPR AHSRWAGLILLLLLEAPILLGVRAQPGVSGPQQRPPOPPPOGGQOQVNGERG ISIPDHGTCQPTISPLCTDIAYNOTIMPNLGHTNSDAGLEVHOFTPLVYKVCASLEL KEFLCSAMTAVCTVLEQALPPFCRSCRAQCEALMNKFGQWPDULKCBKFPVGRALRV ELCVQONTSDKGTPTLPEFTSNFGHGQGGYRGYPGGAGPVERGKFCSPRALRV PSYLNHFELGKDCGAPCEPTKVYGLMVFGEPLFELRSRTVIGWSVLCACASTLFTVLT VLVDMRFSYPERPIIFLSGCYTAVAVYAGLELDRVVCNDKFAEDGARTVAQGT KEGCTILFEMLYFFSMASINWVILSLTWELAGKMGHEALANSOYFHLAWAYPA IKATITLALQGDVGLVGVNGLNVADLGRFVLAFLVFILFIFGFSFLAGVLSLF RIRITMKHDGKTEKLEKLMYRIGVSVLYTPATVIACFYEQAFRFDQWERSWVAQ SCKSYAIFCPHLQGGGYPHPHPSDPFTVMIKYLMTLIVGTSIGFWMVSGKTLNAS RKF			


```
FT S CTVPGSVVYVYGLKRRGELLGLTQTCVRSONVSKASPHTLNSLNCMKINSKLGGINVIL
FT R SPQSLNSEPVLFIGCHLTRSSSLASSDSTSSIAHCDSSIACLVGSMDGHTQFSPTF
FT D TQPRHRTIVDMCEMTREALINERKSTGKPKHIIYRAGIADVTYDEINQTELRAVR
FT F ACAMIEYGFQPGITFIGLDVTHHTRFLFAANEKDRVGNQSNVPAGTLVETGITVNNLFE
FT K YLVSHAGIQGTSRPTXYVVMWDNSIPSAIHMTYQLCHTQSRCTRSVIPSPIVYA
FT L LVAQRAKILMADENFMRPRLCVSNF"
FT CDS complement(join(10723..10767,11212..11658,11707..11823,
FT 12477..12917,13249..13431))
FT /product="T23D8.3"
FT /note="protein predicted using Genefinder"
FT /note="cDNA EST cm16f9 comes from this gene"
FT /note="cDNA EST CEESF93R comes from this gene"
FT /note="cDNA EST CEESF93FC comes from this gene"
FT /db_xref="PID:e304448"
FT /translation="MGKRAFIDKASQSPRLVPDNRSEKFKPTQEHLEEQKYGV
FT Y YDDYDIYLOHMRAVPEMKLENVHEVEKTTIKTSSSAPPAPPGLVGLKALKEEF
FT F DEDVANALEEVTDDRNTGELEDNFITLAGLLDERTTYRSTRGDESEEDDEDD
FT E MYDDYNDDELFGEEAVGEIRVERADQVIDNAPEELMDREYNTDQIGELDGGDDYDVG
FT A LEPNAGRLHLKINDKGPSNAEYDEELAKHYVRERMLIEGVINKDEEYIIVEVDGT
FT N
...
Note: remainder of annotations omitted.
```

```
Query Match 7.6%; Score 52; DB 8; Length 34576;
Best Local Similarity 40.1%; Pred. No. 5.83e-07;
Matches 120; Conservative 44; Mismatches 130; Indels 5; Gaps 5;
```

```
Db 2910 ttcaaatgtcgaataatctgattacaatacaacagattttcccaaatctctcgacata 2969
QY 80 TNCARYTNTCYCAYGGNATGARTYACARAAYATGMGNTNCCNAAVYNTNGNCAYG 139
Db 2970 caacacaaatcgagctggtccagcaattgcgaattcaatccatttaataagtta-aa 3028
QY 140 ARACNATGAARGARG-TNYTNGARCARGCNGCNGTGGATNCCNTNGTNGTNGAARCAR 198
Db 3029 tgcctgaagaatattcgtctcttcttctgactctatcgtacacgtctgtcacag-t-ac 3086
QY 199 TGCAVCCNGAYACNAARARTTITNTGYWSNTNTTGYCNCNGTNGTNGAYGAY 258
Db 3087 t-cgaataacaaatcaacacatgctcgagaattgtgttatctgcgaataatggatgcgag 3145
QY 259 YTNAGYACACNATNCARCCNTGYCAYWSNYTNGYGTNCARGTNAARGAYMGNTGYCN 318
Db 3146 tcattaatgaaaagtgttgatttcaatgccagatacaattggattgttaacaaattccc 3204
QY 319 CCNGTATNGSNGCNTTGYGNTTYCCNTGGCCNGAYATGYTNGARTGYGAYMGNTTCC 377
```

```
RESULT 13
LOCUS CET23D8 34576 bp DNA INV 26-NOV-1996
DEFINITION Caenorhabditis elegans cosmid T23D8.
ACCESSION Z81128
NID g1628237
KEYWORDS
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryoteae; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditidae;
Caenorhabditis.
REFERENCE 1 (bases 1 to 34576)
AUTHORS Wild,A.
Direct Submission
TITLE Submitted (21-OCT-1996) Nematode Sequencing Project, Sanger Centre,
JOURNAL Hinxton, Cambridge CB10 1RQ, England and Department of Genetics,
```

REFERENCE
AUTHORS

Washington University, St. Louis, MO 63110, USA. E-mail: jesus@anger.ac.uk or rwenematode.wustl.edu
2 (bases 1 to 34576)
Willson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M., Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J., Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A., Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M., Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N., Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B., O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A., Saunders,D., Showkeen,R., Smaldon,N., Smith,A., Sonhammer,E., Staden,K., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M., Vaughan,R., Waterston,R., Watson,A., Weinstock,L., Wilkison-Sproat,J. and Wohldman,P.

2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans
Nature 368 (6466), 32-38 (1994)
94150718

Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.

IMPORTANT: This sequence is NOT necessarily the entire insert of clone T23D8. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

The true left end of clone T23D8 is at 1 in this sequence. The true right end of clone T23D8 is at 34576 in this sequence. The true right end of clone T24D1 is at 12749 in this sequence. Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green, ms in preparation), and other available information.

The start of this sequence (1..104) overlaps with the end of sequence CER24D1.

FEATURES
source

1..34576
/organism="Caenorhabditis elegans"
/clone="T23D8"
/chromosome="I"
join(2786..3258,3758..4122,4187..4325,4866..5107,
5220..5340,5688..5906,6332..6479)
/note="protein predicted using Genefinder"

CDS

/codon_start=1
/product="T23D8.g"
/db_xref="PID:e275930"
/db_xref="PID:g1628244"
/translation="MHRHILILFLGCLSDAQRLSSTSISSMNGFSTRKCEHITIPM
CKNLDTNQTVPNLLGHTTQSEAGPAIAQENPLIKVKCEDIRLFTLVYAPVCTVLE
KPTQPCREILCSAKNGCESLMKKEFGWPQDLCNKFPPVTDLCVGRKSSSSKSS
NDVTFGVSTIANEVLSPKKPHMHTTSGSHFSLPLSLGRLEPCLTUCADNQVPM
EDGRVRLIRIWTAAWVACVFCLETLVTLVDLSRFAYVPRPLLYALCYLAISTV
YMLGVGEDFACLTGTPPTTLVTQGGENVGCSALAVHYFFEMSCAWLVLCILAW
IALNLKWAESAIALSPFHMCWGPVAVLSVTLVNSVDGDFVTGICVGNLNP
ALYVFFFTPIVSLAIGAVLLGVCIWMSIRSIKIQHADVERNTSKLKLMLRIGA
FALVYSLPTANNAIIMWYQAVNMPAWLEWHLHRCVRLQDRELFGFTYPPVDDCPMDPK
VAAPEIIVFLKVVSQLVVGITCAIWWVSSKTSYHKAYLALSRSRPTVPAHVQVNV
MR"

CDS

complement(join(7258..7391,7494..7638,8798..8860,
9211..9343,10019..10101,10150..10212))
/note="protein predicted using Genefinder"
/codon_start=1
/product="T23D8.f"
/db_xref="PID:e276248"
/db_xref="PID:g1628243"
/translation="MVEGGTVIVKYLFLANLVWVGGSLIIVSGTILQKFDNVLDI
LGRERLATPILLVIGLCTGLGFCGGAIRNYCLITVFAVILLALITGMVRYHES
RGVESAWDKTHOLFECGVTNNSDLTFTTIPDSCCIEETEGCARENAFLPEPCGHS
VEQWVLKNGAWGGICAVLAAIQLVGCFACCLSKSLKDFHDFY"
complement(join(10723..10767,10978..11142))
/note="protein predicted using Genefinder"
/codon_start=1
/product="T23D8.e"

CDS

H2A; histone H4; Mouse bright protein like; transporter protein.
Caenorhabditis elegans.
ORGANISM
Eukaryotes; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabditia; Rhabditidae; Rhabditina; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 34576)
Wild, A.
Direct Submission
Submitted (21-Oct-1996) Nematode Sequencing Project, Sanger Centre,
Hinxton, Cambridge CB10 1RQ, England and Department of Genetics,
Washington University, St. Louis, MO 63110, USA. E-mail:
jes@sanger.ac.uk or rwnematode.wustl.edu
2 (bases 1 to 34576)
Willson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M.,
Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,
Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A.,
Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,
Johnston, L., Jones, M., Kersey, J., Kirsten, J., Laister, N.,
Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,
O'Callaghan, M., Parsons, J., Percy, C., Rifkin, L., Roopra, A.,
Saunders, D., Showkeen, R., Smalton, N., Smith, A., Sonhammer, E.,
Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,
Vaughan, K., Waterston, R., Watson, A., Weinstock, L.,
Wilkinson-Sproat, J., and Wohlman, P.
2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans
TITLE
NATURE 368 (6466), 32-38 (1994)
JOURNAL
94150718
MEDLINE
COMMENT
Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
IMPORTANT: This sequence is NOT necessarily the entire insert of
clone T23D8. It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.
The true left end of clone T23D8 is at 1 in this sequence. The true
right end of clone T23D8 is at 34576 in this sequence. The true
right end of clone T24D1 is at 12749 in this sequence. Coding
sequences below are predicted from computer analysis, using the
program Genefinder (P. Green, ms in preparation), and other
available information.
The start of this sequence (1..104) overlaps with the end of
sequence CET24D1.
FEATURES
Source
Location/Qualifiers
1..34576
/organism="Caenorhabditis elegans"
/clone="T23D8"
/chromosome="I"
join(2786..3258,3758..4122,4187..4325,4866..5107,
5220..5340,5688..5906,6332..6479)
/note="protein predicted using Genefinder; Similarity to
Drosophila Tissue polarity protein (TR:Q24760)"
/codon_start=1
/product="T23D8.1"
/db_xref="PID:e275930"
/db_xref="PID:g1628244"
/translation="MHRHILFLGCLSDQRSLSSISMSMGFSTRKCEITIPM
CKNLNDNOTVFNPLGHTQSEAGPAIQNFLLKVKCEDIRFLCTVYAPVCTTVE
KPTQPCREGLCKAGNCEFLMKKFGFOWPOLDCNKEPVDLGVGNSSSSSKKSS
NDVFGVSTIANEVULSPKKCPHMHHTSGSHPLSLRPLSCITCADNOVPM
FDRGRILRIWTAANSVACVCSLFTVFLVDLSRFAYVPVRIALAYLAISTV
FLAANKWGAESTAAALSPYFAMCWGPVAVLSVTVNSVDGDFVFTGICVGNLNS
ALVYFETPTVSLAGVLLVCGIWSMIRSVIKIQAQVVERNIKLEKLMRIGA
FATVSLPTANRAIMWYQAVNPANLEGLHRCVRLQDRLELFGFTVDDCPMDPK
VAAPELIVFLKKYQLVVGITCAIWVSSKTSLSYKAYLRELSFPTVADQVN
MR"
complement(join(7258..7391,7494..7638,8798..8860,
9211..9343,10019..10101,10150..10212))
/note="protein predicted using Genefinder; Similarity to

Rat CD63 antigen (SW:P286648)"
/codon_start=1
/product="T23D8.2"
/db_xref="PID:e276248"
/db_xref="PID:g1628243"
/translation="MVEGGVTIVKYLFLANLVLVWVGLSLIIVGSLIQLKFDNVLDI
IGDERATPILLVIGSLCTLLGLGCCGAIRENYCLTVSAFVLLALLIIGVWYHES
RGVESADWKTHQFCEGCVTNSDDWLTFTTTPDSCIEEIEGACARENAPLFEPCGHS
VEQWLVKNGAMVGGICAVLAAIQLVGCFACCLSKSLKDFHDFY"
complement(join(10723..10767,11212..11658,11707..11823,
12477..12917,13249..13431))
/note="protein predicted using Genefinder; cDNA EST cml659
comes from this gene; cDNA EST CEESF93R comes from this
gene; cDNA EST CEESF93C comes from this gene"
/codon_start=1
/product="T23D8.3"
/db_xref="PID:e304448"
/db_xref="PID:g1865750"
/translation="MGKKAFIDKKSQSFRLVDPNRSERFKPTQEHLEEQKYG
YDDVDYLQMRVAVNEPMKLENVHEEVETTKTSGSAFFPAPPFLFGLVGLKKE
FFEDVANALEEVTDDRTGELEDNFTLAGGLDERTTVYRSTRGDEEEDDDDE
DDEWDYNDDELFGGEAVGEIRVERADQVIDNAFELMDREYNTDQIGELDDYD
VGALENAGRLHKLINDKSPSNAEYDEELAKHYRMRMLIEGVIKDEYEIVVE
DEGTNKKMKWDCSEFATQYTIYNHPTLIREPGLSKALKRDKFAVEEMDAEDED
DDEMDDDDDSEFSTVTFPRKNETPEQSRKRAKLRKRRKANKKMTFAE
EKRLAKRGIGQIKARPI"
complement(join(13883..14240,14292..15409,15456..16341,
16423..16556,16705..16807,17668..17721,17774..18278,
18330..18656,18888..19113,19170..19705,19751..19828,
19884..19991,20041..20214))
/note="Similarity to Yeast nuclear transport protein
(SW:P32497); cDNA EST yk44d4.5 comes from this gene; cDNA
EST yk33b1.5 comes from this gene; cDNA EST yk37d3.5 comes
from this gene; cDNA EST yk8e5.5 comes from this gene;
cDNA EST yk8e7.5 comes from this gene; cDNA EST yk37b1.5
comes from this gene; cDNA EST yk4b7.5 comes from this
gene; cDNA EST yk30g3.5 comes from this gene; cDNA EST
yk8e5.3 comes from this gene; cDNA EST yk8e7.3 comes from
this gene; cDNA EST yk44d4.3 comes from this gene; cDNA
EST yk37b1.3 comes from this gene; cDNA EST yk33b1.3 comes
from this gene; cDNA EST yk37d3.3 comes from this gene"
/codon_start=1
/product="T23D8.4"
/db_xref="PID:e304461"
/db_xref="PID:g1865751"
/translation="MSYHQDDRMSLSVNTTMTNQFSDQRCYSSNGSTCENVPEN
MQMNTQAAEQAIRLWFNTPDPQRLMAKTIWIRQDKRFAQVDQANMPCVQQLN
IYDGLKPPQVLPISYYAQLWYNLLDIRRFTPLIISPIHQVQVQFCPRENGPQD
FRELICNLISLNQKDPKMKCANQVQIFNCIIMGYKNEKLRTEFAQHLKFLKLVGT
LSEVFNQVHPGMINPAIFIIFRIISKDTLKDYFWNNNPHQPPPTGLIILKLA
VMIGSVELIAGQNPETLPQNELAHLIQLIIRTFDGLGLLHSDAIDGFFVRSGVGA
ITVVQYPNDLIRAGCKLLQVSDAKALAKTPLENLIPELLRLIEIHPDDEVISGT
GFSUNVAHQHVKDIAIRSNALFLLHIIISKYPRDLDELDAKRNRCVEICNCLRT
LNFLNFIPTPNGETAGNEQVQKFEIIDLKMLSCSGEDMDPTGLLELRS
TILRSFTLLRTPVPKDGVDNIDENKLENLIGHICAAYSWFRQNPNTQSTKQ
LVERTISLLVLMEOCGAEVAQYSYIDCLPLNLNGNQVQPTFIHNLVVCOKILE
HCTRADITIDRPMLEGLTNRNSDIAKAANSLLSRFPENDLLAASRPFHAKEDSD
SDTSSSEDEVEDQVNSAKAFRODDFMAGPEDEKRVRAQKDKKDELDYGLIKQNR
DAKSNLDNLLTGFLSKAYDKSTVFORQNVANPRFYIRSLVEILEDYKLNWQDK
EVSALKSNNAKALPPLRLKRYIKDQQLDLYTVRNPDEGDEYETPDEDDDDFG
EVESKAEKSPGKSEKAAVSDSDSDSDSDSNSSSEPSNDDSDSVTKMQLLR
YFLKKEFRSVDKDKKFKKRVIRKVEAEDDDADWTVPVNRKSVHFDPEKEVTHD
VAKKLNVSASRGKRTTDRNQHVANRLKLEVESEKELGLGINVKISFCILSALFEL
NAKIDHMEYETFTMTLTQVNSLDDLLIGDVRKLSVYAEEDENADQTOEYRIGGS
ILLAVQVMDIELAKIQNADCHNDYIEKLAKDMCSLIEKAEKVELRNDSDYFYL
HEVCYVLMIEHAYKYQDNEDDAGKMLCNKTYITLDEKRLQAMLCRHFYK
AVHDKWRARLLIMSHMOALVDHSDVDTOILYNTICQLGICAFRGFTREAHQGLS
EIQNTQAKELLQAQVTRQHEKTAQEKIDRSQVQVYHHIHVELMECYVLICSMILL
EIHPMAECEEMRRMRLSRFHYLQKSEKASLTGPPENTREHNVHVAASKAMLDGWRK
CQDIYVNDKMNQVKNMFLFNAETVKGMVVRIOBEESLRTYLLTYATVYSLKKLAD
LFELSKKDVHSIISKMIQBELSATLDEPTDCLIMHVRPESRLQMLAINSLDKLQTLA
ENNEQILERTGRGGYOGPGSWFPGRNERQDKQKGGGGYQGERRRGGQDGRKNGWG
SQSGQQRHHPKPRAF"

CDS

CDS

CDS

CDS



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by Intelligenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Dec 9 08:43:36 1997; MasPar time 1043.08 Seconds
Tabular output not generated. 997.528 Million cell updates/sec

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCCTTCATGGCCT.....AAAAAAGAGCGCGCGC 2027
Comp: CTTAAGCGGAAGTACCGGA.....TTTTTTTTTTTCGCGCGCG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 707517 seqs, 256659390 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

EST-STS

1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99

EST-STS-TWO

100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
170:EST170 171:EST171 172:EST172 173:EST173 174:EST174

Statistics: Mean 12.162; Variance 2.824; scale 4.307

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	473	23.3	480	134	yy6409.r1 Homo sapie	0.00e+00
2	453	22.3	460	173	z270811.s1 Soares fet	0.00e+00
c 3	416	20.5	446	101	yy6409.s1 Homo sapie	0.00e+00
4	413	20.4	462	174	zd70611.r1 Soares fet	0.00e+00
c 5	408	20.1	445	92	yy83c06.s1 Homo sapie	0.00e+00
c 6	363	17.9	433	156	ze76g10.s1 Soares fet	0.00e+00
7	359	17.7	426	72	ys74d07.r1 Homo sapie	0.00e+00
8	337	16.6	406	61	yl28c05.r1 Homo sapie	0.00e+00
9	331	16.3	340	48	Human aorta CDNA 5'-e	0.00e+00
10	328	16.2	396	178	mm97f07.r1 Stratagene	0.00e+00
11	329	16.2	414	156	ze76g10.r1 Soares fet	0.00e+00
12	322	15.9	434	45	yy73a07.r1 Homo sapie	0.00e+00
c 13	321	15.8	469	61	yl28c05.s1 Homo sapie	0.00e+00
c 14	319	15.7	326	92	yy83e07.s1 Homo sapie	0.00e+00
15	316	15.6	330	48	Human aorta CDNA 5'-e	0.00e+00
16	313	15.4	344	99	yy83c06.r1 Homo sapie	0.00e+00
17	299	14.8	301	110	Human aorta CDNA 5'-e	0.00e+00
18	299	14.8	313	48	Human aorta CDNA 5'-e	0.00e+00
19	298	14.7	303	49	Human aorta CDNA 5'-e	0.00e+00
20	294	14.5	300	48	Human aorta CDNA 5'-e	0.00e+00
21	290	14.3	293	48	Human aorta CDNA 5'-e	0.00e+00
22	290	14.3	293	48	Human aorta CDNA 5'-e	0.00e+00
23	290	14.3	344	61	yl26d03.r1 Homo sapie	0.00e+00
24	287	14.2	299	48	Human aorta CDNA 5'-e	0.00e+00
25	285	14.1	287	48	Human aorta CDNA 5'-e	0.00e+00
26	279	13.8	281	110	Human aorta CDNA 5'-e	0.00e+00
27	273	13.5	273	48	Human aorta CDNA 5'-e	0.00e+00
28	269	13.3	283	164	Human aorta CDNA 5'-e	0.00e+00
29	256	12.6	263	110	Human aorta CDNA 5'-e	0.00e+00
30	254	12.5	266	48	Human aorta CDNA 5'-e	0.00e+00
31	251	12.4	300	49	Human aorta CDNA 5'-e	0.00e+00
32	241	11.9	241	48	Human aorta CDNA 5'-e	0.00e+00
33	230	11.3	247	110	Human aorta CDNA 5'-e	0.00e+00
c 34	226	11.1	227	61	Human aorta CDNA 5'-e	0.00e+00
35	213	10.5	215	164	Human aorta CDNA 5'-e	0.00e+00
c 36	212	10.5	222	73	Human aorta CDNA 5'-e	0.00e+00
c 37	210	10.4	216	48	Human aorta CDNA 5'-e	0.00e+00
38	209	10.3	299	148	Human aorta CDNA 5'-e	0.00e+00
39	200	9.9	201	140	Human aorta CDNA 5'-e	0.00e+00
40	200	9.9	231	65	Human aorta CDNA 5'-e	0.00e+00
41	200	9.9	258	98	Human aorta CDNA 5'-e	0.00e+00
42	194	9.6	198	182	Human aorta CDNA 5'-e	0.00e+00
43	193	9.5	195	182	Human aorta CDNA 5'-e	0.00e+00
c 44	192	9.5	204	65	Human aorta CDNA 5'-e	0.00e+00
45	184	9.1	210	48	Human aorta CDNA 5'-e	0.00e+00

ALIGNMENTS

1 N93894 480 bp mRNA EST 05-APR-1996
LOCUS yy64d09.r1 Homo sapiens CDNA clone 278321 5'
DEFINITION N93894
ACCESSION N93894
NID 91266203
KEYWORDS EST.
SOURCE human clone-278321 primer-mob.REGA+ET library-Soares multiple


```

Db 61 tattttacaacttttaagtaaaactactatgtattacaggttaagctacaatgggtt 120
|||||
Cp 1926 TATTTTACAACTTTTAAAGGTAACACTACTATGTATATTACAGGTAAGCTACAATGGGTT 1867
|||||
Db 121 taatttgcacaaagtaagtaagaatgttttaaaacagcgttaaaagtactcaagtcatt 180
|||||
Cp 1866 TAAATTCGAAAAGTTAAGTAGAAGATGTTTAAACAAGGCTTAAAGTACTCAAGTCAATT 1807
|||||
Db 181 ataaaaattatatcttttgccttttacttgaagaatcatctatagaaaatggttaagt 240
|||||
Cp 1806 ATAAATTTATATCTTTTGCCTTTTACTTGAAGAATCATGCTATAGAAATGGTTAAATGT 1747
|||||
Db 241 gcttctaataaatggaagtattagctggaatgtgacatgtgaacaggtttaagttccc 300
|||||
Cp 1746 GCTTCTTAATAATGGAAGTATTGAGCTGGAATGTGATACATGTAACAGTTTAAAGTTCCC 1687
|||||
Db 301 attgaaggtatacaaatgatgaatcgttctgttaagcacttagacactgagctcagctcgag 360
|||||
Cp 1686 ATTGAAGGTATAAATGATGAATT-GTTGTAAAGACTTAGACACGAGCTCAGTCTGGAG 1628
|||||
Db 361 ctgatgaagatgttgagataaacagccagctttatctcaacaggggttctgacccacaagt 420
|||||
Cp 1627 CTGATGAAGATGTTGAGATAACAGCAGCAGCTTATCTCAACAGGGTTTGTGACCCACAAGT 1568
|||||
Db 421 ttggggcacagagaaaaattggaagcaatttgcattgtatga 460
|||||
Cp 1567 TTGGGCGACAGAGAAAATTGAAGCAATTTGCGATGTTATGA 1528
|||||

RESULT 3
LOCUS N64004 446 bp mRNA EST 01-MAR-1996
DEFINITION Y764009.s1 Homo sapiens cDNA clone 278321 3'.
ACCESSION N64004
NID g1211833
KEYWORDS EST.
SOURCE human clones-278321 primer-m13 -40 forward library-Soares multiple
sclerosis 2bNHMSP vector-pf73D (Pharmacia) with a modified
polylinker V-type: Phagemid host-DHI0B (ampicillin resistant)
Rsai1-Not I Rsite2-Eco RI 46 year old male. 1st strand cDNA was
primed with a Not I - oligo(GT) primer
[5'-GTGTACCAATCTGAGTGGGAGCGCGCATTTTATTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified p77 vector (Pharmacia). Library went
through one round of normalization to a Cot = 5. Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4
multiple sclerosis lesions from one patient was kindly provided by
Dr. Kevin G. Becker (NINDS/NIH).
Homo sapiens
ORGANISM Homo sapiens
Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 446)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)

REFERENCE
AUTHORS
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)

TITLE
JOURNAL
COMMENT
Contact: Wilton RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
High quality sequence stops: 225
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

```

```

FEATURES
Source Location/Qualifiers
mRNA 1..446
BASE COUNT 157 a 51 c 87 g 149 t 2 others
ORIGIN
Query Match 20.5%; Score 416; DB 101; Length 446;
Best Local Similarity 98.9%; Pred. No. 0.00e+00;
Matches 440; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

Db 2 taggtataaacagagtgtaaaagtttatatacaagaatataatgttttatctgaatatttac 61
|||||
Cp 2003 TAGGTAAAACAGAGATGTAAGTTTATATACAAGAATATATGTTTATCTCGAAATATTTC 1944
|||||
Db 62 gctgtgttgaagcaatatttttacaacttttaaacgttttaaggttaaaactactatgtattacag 121
|||||
Cp 1943 AGTGTGTGGTTAAGCAATATTTTACAACTTTTAAAGGTAAGTACTATGTATATTACAG 1884
|||||
Db 122 gtaagctacaatgggtttaaatttgcaaaagttaagttaagtaaaatgtttttaaacaaggctta 181
|||||
Cp 1883 GTAAGCTACAATGGGTAAATTGCAAAAGTTAAGTAAGAATGTTTAAACAAGGCTTA 1824
|||||
Db 182 aagttactcaagtcattataataaatttatctcttttgccttttacttgaagaatcatgct 241
|||||
Cp 1823 AAGTACTCAAGTCAATATATAAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCT 1764
|||||
Db 242 atagaatgggttaagtgtcttctaataaagtgaagtaattgttagctggaatgtgatacatg 301
|||||
Cp 1763 ATAGAATGGGTAAATGTGCTTCTTAATAAGTAAGTATTGTAGTGAATGTGATACATG 1704
|||||
Db 302 taacagtttaagttccattccattgaaggtataaaatngatgaattgttgaagacccttagaca 361
|||||
Cp 1703 TAACAGTTTAAGTTCCATTGAGGTATATAAAT-GATGAATGTTGTAGAC-TTAGACA 1646
|||||
Db 362 ctggagttcagctggtgagctgtagaagatgttgagataaacagccagcttttatctcaac 421
|||||
Cp 1645 CTG-AGTCTCAGCTGG-AGCTGATGAAGATGTGAGATAACAGCCAGCTTTTATCTCAAC 1588
|||||
Db 422 aggggttctgaccncaagttggg 446
|||||
Cp 1587 AGGGTGTGTGCCCAAGGTTTGGG 1563
|||||

RESULT 4
LOCUS W77968 462 bp mRNA EST 17-OCT-1996
DEFINITION zd70e11.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
346028 5'.
ACCESSION W77968
NID g1388502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 462)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)

TITLE
JOURNAL
COMMENT
Contact: Wilton RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LLNL ; contact the

```

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1550 Std Error: 0.00

Seq primer: mob.REGA+EF

High quality sequence stop: 424.

Location/Qualifiers

FEATURES

```

source
1..462
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19W."
/clone="345028"
/clone_lib="Soares fetal heart NbHL19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1..>462
BASE COUNT      128 a  133 c  106 g   91 t    4 others
ORIGIN
Query Match      20.4%; Score 413; DB 174; Length 462;
Best Local Similarity 98.0%; Pred. No. 0.00e+00;
Matches 449; Conservative 0; Mismatches 3; Indels 6; Gaps 6;
Db 6  tgcgtcttcgccccgcgtctgcctgcgtacccagcagaccatccagccactgcactcg 65
Qy 544  TGCCTCTTCGCCCGCTCGCTCGATGACCTAGACGAGACCATCCAGCCATCCCACTCG 603
Db 66  ctctgn-tgcagtggaagcagcgtgcgcgcgcgtatgtccgccttcgn-ttccctg 123
Qy 604  CTCTCGGTGCAGGTGAAGCAGCGCTCGCCCGGTCTATGTCCGCCTTCGCCCTGG 663
Db 124  ccgcacatgcttgagtcgacgcttccccccaggacacgaccttgcctccccctccgc 183
Qy 664  CCCGACATGCTTGAGTGCACCGCTTTCCTCCCGAGGACACGACCTTTCATCCCTC-GC 722
Db 184  tagcagn-accacctctctccagccaccaggaggaagctccaaaggtatgtgaagcctgcaa 242
Qy 723  TAGCAGCGACCACTCTCTCCGACCCACCAGGAGGCTCCAAAGGTATGTGAAGCCTGCAA 782
Db 243  aaataaaatgatgatgacaacagacataatgaaacgctttgtaaataatgatttgcact 302
Qy 783  AAATAAAATGATGATGACACGACATAATGGAACGCTTTGTAAATAATGATTTGCACT 842
Db 303  gaaaataaagtgaaggagatacctcacatcaaccagacagataccaaatcatcctggaaga 362
Qy 843  GAAATAAAGTGAAGGAGATACCTTACATCAACCCGAGATACCAAAATCATCTCTGGA-GA 901
Db 363  ccaegagaagaccatttcaagctgaacggtgtgtccgaaaggagaccttgaagaatcg 422
Qy 902  CCAAGAGCAAGACCATTTACAGCTGAACGGTGTGTGCCGAAGGGACCT-GAAGAAATCG 960
Db 423  gtgctgtggctcaagacagcgttgcactgtga 460
Qy 961  GTGCTGTGGCTCAAGACACGCTTGCAGTGCACCTGTGA 998

```

```

RESULT 5
LOCUS      N32415      445 bp      mRNA      EST      10-JAN-1996
DEFINITION Yw83c06.s1 Homo sapiens cDNA clone 258826 3'.
ACCESSION  N32415
NID         g1152814
KEYWORDS   EST.
SOURCE     human clone=258826 primer=n13 -40 forward library=Soares placenta
           8to9weeks 2NBHP8to9W vector=pt7T3D (Pharmacia) with a modified

```

polylinker host=DH10B (ampicillin resistant) Rsite1-Not I
Rsite2-Eco RI two placentae: one from 8 weeks and another from 9
weeks post conception. 1st strand cDNA was primed with a Not I -
oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7T3 vector (Pharmacia). Library
constructed by Bento Soares and M. Fatima Bonaldo.

ORGANISM

Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 445)
Hillier, L., Clark, N., Dubuque, T., Eliston, K., Hawkins, M.,
Hollman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)

TITLE JOURNAL COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
High quality sequence stops: 314
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

```

Source
1..445
/organism="Homo sapiens"
/clone="258826"
<1..>445
BASE COUNT      156 a  55 c   87 g   147 t
ORIGIN
Query Match      20.1%; Score 408; DB 92; Length 445;
Best Local Similarity 98.6%; Pred. No. 0.00e+00;
Matches 434; Conservative 0; Mismatches 2; Indels 4; Gaps 4;
Db 1  aaacagggatgaacgtttatatacaagaataataagtttattctgaaatttaccagt 60
Cp 1998  AAAACAGGATGTAAGTGTATATACAGGAATATAATGTTTATCTGAAATATTTACAGTGT 1939
Db 61  tggtaaagcaatatttttacaacttttaaggttaagtaagaatgttttaacaagcgttaagta 180
Cp 1938  TGGTTAAGCAATATTTTACAACTTTTAAAGGTAACACTACTATGATATATATACAGGTAAG 1879
Db 121  ctacaatgggttttaatttcgaaagttaagtaagaatgttttaacaagcgttaagta 180
Cp 1878  CTACAATGGGTTTAAATTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCTTAAAGTA 1819
Db 181  ctcaatgcaattataaaatttatattttgoccttttacttggaagaatacatcctataga 240
Cp 1818  CTCAGTCAATTAATAAATTTATATCTTTTGCCTTTTACTTTGAAGAAATCATGCTATAGA 1759
Db 241  aatgggttaagtgtctctaataaagtgaagtattgtagctggaatgtgacatgtaaca 300
Cp 1758  AATGGTTAATGTGCTTCTAATAATGAAGTATGTAGCTGGAAATGTGATACATGTAAACA 1699
Db 301  gtttaagtcccatggaaggtataaaatgatgaattgttgtaagacttagacactgaggt 360
Cp 1698  GTTTAGTTCCTCATTGAAGGTATAAATGATGAATTTGTTGTAGACTTAGACACTG-AGT 1640
Db 361  ctcaagtctggagctgatgaagatgttgagataacacccagcagctttatcccccaacaggggt 420
Cp 1639  CTCAGTCTGG-AGCTGATGAAGATGTTGAGATAACAGCCAGCTTTTATCTC-AACAGGG-T 1583

```

Db	421	ttgtgacccacaagtgttg 440
Cp	1582	TTGTGACCCACAAGTTGGG 1563
RESULT	6	AA024772 433 bp EST 13-AUG-1996
LOCUS		zef6q10.s1 Soares fetal heart NBHH19W Homo sapiens cDNA clone
DEFINITION		364962 3'.
ACCESSION		AA024772
NID		g1489678
KEYWORDS		EST.
SOURCE		human.
ORGANISM		Homo sapiens
REFERENCE		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 433)
AUTHORS		Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE		The WashU-Merck EST Project
JOURNAL		Unpublished (1995)
COMMENT		Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Location/Qualifiers 1..433 /organism="Homo sapiens" /note="Organ: heart; Vector: pT7n3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGCGGCCGCATCTTTTTTTTTTTT 3'] , double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7n3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W." /clone="364962" /clone_lib="Soares fetal heart NBHH19W" /sex="unknown" /dev_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" complement(<1..>433) 150 a 51 c 141 t 4 others
BASE COUNT		150 a 51 c 141 t 4 others
ORIGIN		
Query Match		17.9%; Score 363; DB 156; Length 433;
Best Local Similarity		99.2%; Pred. No. 0.00e+00;
Matches		369; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Db	1	gtaaacaggagtgtaagtttatatacaagaataataattgttctgaattattcaggt 60
Cp	2000	GTA AACAGGATGTA AAGTTTATATA CAAGAATAA TATGTTTATCTGAATATTACAGT 1941
Db	61	gttggttaaaccaattttttcacacttttaaaaggtaaactactatgtattaccaggt 120
Cp	1940	GTTGGTTAAAGCAATATTTTACAACTTTTAAAGGTA AAACACTATGTATATTACAGGTA 1881

Japan
Phone: 0886-65-2888
Fax : 0886-37-1035.

FEATURES
source

Location/Qualifiers
1..340
/organism="Homo sapiens"
/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"
BASE COUNT 113 a 64 c 38 g 122 t
ORIGIN

Query Match 16.3%; Score 331; DB 48; Length 340;
Best Local Similarity 98.5%; Pred. No. 0.00e+00;
Matches 335; Conservative 3; Mismatches 1; Indels 1; Gaps 1;
Db 1 catcttcacgctccagactgagactcagtgcttaagtcttaacaacaattcatcttt 60
QY 1616 CATCTTCATCAGCTCCAGACTCAGACTCAGTGTCTAAGTCTTACAACAATTCATCATTT 1675
Db 61 ataccctcaatgggaactaaactgttacatgtatcacattccagctacaacttccat 120
QY 1676 ATACCTTCAATGGGAACCTAAACTGTATCATGTATCACATTCCAGCTACAATACTTCCAT 1735
Db 121 ttattagggcacattaacatttctatagcatgattctcaagtataaagcaaaagat 180
QY 1736 TTATTAGAGCACATTACCACTTCTATAGCATGATTTCTCAAGTAAAGGCAAAAGAT 1795
Db 181 ataaatttataatgacttgactcttaagcctgttttaaacatttcttacttaact 240
QY 1796 ATAAATTTTATAATGACTTGTAGTACTTTTAAGCCTGTTTAAACATTTCCTACTTAACT 1855
Db 241 ttgggcaataaaccattgtagcttacctgtaataatatacatagtrgrttacotttaaaa 300
QY 1856 TTGTG-CAAAATTAAGCCCATTTGAGTGTACTGTAAATATACATAGTAGTTTACCTTTAAAA 1914
Db 301 gttgtaaaartattctttracacacactgtataatttc 340
QY 1915 GTTGTAAAATATTGCTTTAACCAACACTGTAAATTTTC 1954

RESULT 10
LOCUS AA073862 396 bp mRNA EST 04-OCT-1996
DEFINITION mm97f07.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
36389 5' similar to TR:G1151260 G1151260 TRANSMEMBRANE RECEPTOR.

ACCESSION AA073862
NID g1595602
KEYWORDS EST.

SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.

REFERENCE 1 (bases 1 to 396)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)

TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:323325
Seq primer: primer name ambiguous
High quality sequence stop: 390.

FEATURES
source

Location/Qualifiers
1..396
/organism="Mus musculus"
/strain="NIH/Swiss"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. 93
pooled NIH/Swiss 13 day embryo hearts. Average insert
size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
GAATTCGGCAGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3'"
/clone="536389"
/clone_lib="Stratagene mouse heart (#937316)"
/sex="pooled"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
BASE COUNT 93 a 126 c 103 g 74 t
ORIGIN

Query Match 16.2%; Score 328; DB 178; Length 396;
Best Local Similarity 92.5%; Pred. No. 0.00e+00;
Matches 368; Conservative 0; Mismatches 28; Indels 2; Gaps 2;
Db 1 agcagcgggcgct-ggattccgctggtcatgaagcagtgccaccggacaccaaagt 59
QY 476 AGCAGGCGCGGCTTGGATCCCGTGGTCATGAAGCAGTGCCACCGGACACCAAGAAT 535
Db 60 tctgtgctgctcttcgccccctgtctgtctgcagcactagatgacacatccagcgt 119
QY 536 TCGTGTCTCGCTCTTCGCCCGCTGCTGCTGATGACCTAGACGAGACCATCCAGGCAT 595
Db 120 gtacactcgctctg-gtgcaggtgaagcagcgctgcgcccgggtcatgtccgctctcgct 178
QY 596 GCCACTCGCTCTCGGTGCAGGTGAAGGACCGCTGCGCCCGGTCATGTCGCCCTTCGGCT 655
Db 179 tccccggccagacatgctggagtgcgaccttcccgccagggacaaagacacctgcctcc 238
QY 656 TCCCTTGGCCCGACATGCTTGAGTGGCAGCGTTTCCCGCAGGACACGACCTTTGCATCC 715
Db 239 cctcgtagtagcagaccactcctcggcgccacagaggaaagctcccaagtggtgtaag 298
QY 716 CCTCTGCTAGCAGCGACCCACTCTCTGCCACCGCAGGAAGCTCCAAAGTATGTGAAG 775
Db 299 cctgcaaaacaaagatgagcagcacaacgacatcatggaacccctttgtaaaatgact 358
QY 776 CTGCAAAATAAATAATGATGATGACACACACATATATGAAACGCTTTGTAATAATGATT 835
Db 359 tcgactgaaaatcaaaagtgaaggagataaacgtacatc 396
QY 836 TTGCACCTGAAATAAAGTGAAGGAGATACCTACATC 873

RESULT 11
LOCUS AA024771 414 bp mRNA EST 13-AUG-1996
DEFINITION ze76g10.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
364962 5'.
ACCESSION AA024771
NID g1489677
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 414)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevisakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilton,R.
The WashU-Merck EST Project
Unpublished (1995)
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilton RK

WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 288.

FEATURES

Location/Qualifiers

source

1..414

/organism="Homo sapiens"

/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bernaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHL19W."

/clone="364962"

/clone_lib="Soares fetal heart NBHL19W"

/sex="unknown"

/dev_stage="19 weeks"

/lab_host="DH10B (ampicillin resistant)"

<1..>414

107 a 123 c 95 g 85 t 4 others

mRNA

BASE COUNT

ORIGIN

Query Match 16.2%; Score 329; DB 156; Length 414;
Best Local Similarity 97.8%; Pred. No. 0.00e+00;
Matches 360; Conservative 0; Mismatches 3; Indels 5; Gaps 5;

Db 1 tgaagcagtcgacccggacaccaaagaattcctgtgctgctcttcgccccgcgtcgcc 60

QY 506 TGAAGCAGTGCACCCCGGACCAAGAGTTCTGTGCTGCTCTGCCCCGCTCGCC 565

Db 61 tcgtagcctagcagacccatccagccatcgccactgctctcgctgcaggtgaagacc 120

QY 566 TCGTAGCCTAGCAGACCATCCAGCCATCGCCTCGCTCGTGCAGGTGAAGACC 625

Db 121 gctgccccggtcatgtccgcttcgntccctgccccggcggaaatgcttgatggac 180

QY 626 GCTGCCCCCGGTCATGTCGCGCTTCGGCTTCGCCCTGGCCCCG-ACATGCTGAGTGGCAG 684

Db 181 cgttccccagacacgacctttgcatccccctgctgtagcagcgaccacctctcgcca 240

QY 685 CGTTTCCCCAGGACACGACCTTTGCATCCCCCTCGCTAGCAGCGACCTCCTGCCA 744

Db 241 gccacggaggaagctccaaaggtatgtgaagcctgcaaaataaaaaatgatgatgacaac 300

QY 745 GCCACCGAGGAGCTCAAAGGTATGTGAAGCCTGCAGCAATAAATAATGATGATGACAAC 804

Db 301 gacattaatggaacgctttgtataaatgatatttgacttgaaattgaaaggaagga 360

QY 805 GACAT-AATGGAAACGCTTTGTAAAAATGATTTTGGCACT-GAAAAAT-AAAAGTGAA-GGA 860

Db 361 gataacct 368

QY 861 GATAACCT 868

RESULT 12

LOCUS

DEFINITION

ACCESSION

NID

KEYWORDS

H44092 434 bp mRNA EST 31-JUL-1995

Y073a07.r1 Homo sapiens cDNA clone 183540 5'

H44092

9920144

EST.

SOURCE

human clone-183540 library-Soares breast 3nBHst vector-pT7T3D
(Pharmacia) with a modified polylinker host-DH10B (ampicillin
resistant) primer-M13RP1 Rsite1-Not I Rsite2-Eco RI Adult human.
1st strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adapters (Pharmacia),
digested with Not I and cloned into the Not I and Eco RI sites of a
modified pT7T3 vector (Pharmacia). Library went through one round
of normalization to a Cot = 20. Library constructed by Bento Soares
and M.Fatima Bernaldo.

ORGANISM

Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 434)

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

TITLE

The WashU-Merck EST Project

JOURNAL

Unpublished (1995)

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

High quality sequence stops: 123

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

Location/Qualifiers

source

1..434

/organism="Homo sapiens"

/clone="183540"

124 a 87 c 72 g 146 t 5 others

BASE COUNT

ORIGIN

Query Match 15.9%; Score 322; DB 46; Length 434;

Best Local Similarity 93.5%; Pred. No. 0.00e+00;

Matches 405; Conservative 0; Mismatches 16; Indels 12; Gaps 12;

Db 2 gctgattttcacttccacitgaggttgcataacatgcataatgcttcaatttctctgt 61

QY 1501 GCTGATTTTCACTTCCACCTGAGGTGTGCTAATCAATGCAATTCCTTCTCTGT 1560

Db 62 ggcccaactgtgggtcacaaccctgttgagataaagctggctgttatcacaactct 121

QY 1561 GGCCCAACTGTGGGTGTCACAAACCTGTGAGATAAAGCTGGCTGTATCTCAACATCT 1620

Db 122 tcacagctccagactgagactcagtgcttaagtcttaccacaatnctatttacc 181

QY 1621 TCATCAGCTCCAGACTGAGACTCAGTGTCTAAGTCTTACAACTTTCATCATTTATACC 1680

Db 182 ttaattgggaacttaactgttacatgtatcacattccagctacaatacttccatttatt 241

QY 1681 TTCAATGGGAACCTAAACTGTTACATGTATCACATTCAGATTAATCTCCATTTATT 1740

Db 242 aggaagcacattaaccatttctatagcagtgatttctcaaggtaaaaggcgaaggta 301

QY 1741 AG-AAGCACATTACCAATTTCTATAG-CACTGATTTCTTAAG-TAAAGG-CAAAAGATA 1796

Db 302 ttaatttttaattggactgaggtactttaaggcctgttttaaaacatttcttact 361

QY 1797 T-AAATTTTAAATG-ACTTGAG-TACTTT-AAG-CCTGTGT-AAACATTTCTTACT 1850

Db 362 taacttggcaaataaacccctgtnggttaccgccgtantttacngtgggtttacct 421

QY 1851 TAACCTTTGCAATTAACCCATTGTAG-CTTACCTG-TAATATACATAGTAGTTTACCT 1908

```

Db 422 ttaaaagtggttaa 434
|||||
Qy 1909 TTAAGAGTTGTAA 1921

RESULT 13
LOCUS H15818 469 bp mRNA EST 27-JUN-1995
DEFINITION Y128c05.s1 Homo sapiens cDNA clone 159560 3'.
ACCESSION H15818
NID 9880638
KEYWORDS EST.
SOURCE human clone=159560 library=Soares breast 3NBHBst vector=PT73D
(Pharmacia) with a modified polylinker host=DH10B (ampicillin
resistant) primer=Promega -2lml3 Rsite1-Not I Rsite2=Eco RI Adult
human. 1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),
digested with Not I and cloned into the Not I and Eco RI sites of a
modified pT73 vector (Pharmacia). Library went through one round
of normalization to a Cot = 20. Library constructed by Bento Soares
and M.Fatima Bonaldo.

ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 469)
Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 340
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source
Location/Qualifiers
1..469
/organism="Homo sapiens"
/clone="159560"

BASE COUNT 144 a 65 c 106 g 150 t 4 others
ORIGIN

Query Match 15.8%; Score 321; DB 61; Length 469;
Best Local Similarity 97.1%; Pred. No. 0.00e+00;
Matches 370; Conservative 0; Mismatches 3; Indels 8; Gaps 8;

Db 1 aggttaaacagagatttaagtttatatacaagaataataatgtttatctgaatattaca 60
|||||
Cp 2002 AGGTAAACAGAGATTAAAGTTTATATACAAAGATATATGTTTATCTGAAATATTTTACA 1943
|||||

Db 61 gtgttggttaagcaaatattttacaacttttaagggttaaacactatgtattattacagg 120
|||||
Cp 1942 GTGTGTTGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATATATATTACAGG 1883
|||||

Db 121 taagtacaatgggttttaatttgcaaaagttaagtaagaataatgttttaacaaggtttaa 180
|||||
Cp 1882 TAAGCTACAATGGGTTTAAATTGCAAAAGTTAAGTAAGAAGTGTTTTAAACAAGGCTTAA 1823
|||||

Db 181 agtactcaagtcgaattataaaatttatctcttttgccttttacttggaagaatcatgcta 240
|||||
Cp 1822 AGTACTCAAGTCAATTATATAAAATTTATATCTTTTGCTTTTACTTTGAAGAAATCATGCTA 1763
|||||

```

```

Db 241 tagaaatggttaatgctcttaataaaagggaagtattgttagcgggaatgtgataca 300
|||||
Cp 1762 TAGAAATGGTTAATGTCTTCTAATAAATGG-AAGTATTGTAG-CTGG-AATGTGATACA 1706
|||||

Db 301 tgtaacagtttaaggttcccatgaagggtataaaatgatgaattgttgaagnccttag 360
|||||
Cp 1705 TGTAACAGTTTAAG-TTCCCATTTGAAGG-TATAAATGAATGATTTGTTGAAG-ACTTAG 1649
|||||

Db 361 gacactgaggtctcagtcngg 381
|||||
Cp 1648 -ACACTGAG-TCTCAGTCTGG 1630

RESULT 14
LOCUS N32424 326 bp mRNA EST 10-JAN-1996
DEFINITION YW83e07.s1 Homo sapiens cDNA clone 258852 3'.
ACCESSION N32424
NID g1152823
KEYWORDS EST.
SOURCE human clone=258852 primer=ml3 -40 forward library=Soares placenta
8to9weeks 2NBHP8to9W vector=PT73D (Pharmacia) with a modified
polylinker host=DH10B (ampicillin resistant) Rsite1=Not I
Rsite2=Eco RI two placentae: one from 8 weeks and another from 9
weeks post conception. 1st strand cDNA was primed with a Not I -
oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT73 vector (Pharmacia). Library
constructed by Bento Soares and M. Fatima Bonaldo.

ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 326)
Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 312
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source
Location/Qualifiers
1..326
/organism="Homo sapiens"
/clone="258852"
<1..>326

BASE COUNT 123 a 33 c 54 g 116 t
ORIGIN

Query Match 15.7%; Score 319; DB 92; Length 326;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 325; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 taaacacagatgtaagtttatatacaagaataataatgtttatctgaatattacagtg 60
|||||
Cp 1999 TAAACAGGATGTAAGTTTATATACAAAGATATATGTTTATCTGAAATATTTACAGTG 1940
|||||

Db 61 ttgggttaagcaaatatttttacaacttttaagggttaaacactactatgtattattacaggtaa 120
|||||

```



```
|||||
Cp 1939 TTGGTTAAAGCAATATTTTACAACTTTTAAAGGTAACCTACTATCTATATTACAGTAA 1880
|||||
Db 121 gctacaattgggttttaattgcacaaagttaagtaagaataatgttttaacaagggtttaaagt 180
|||||
Cp 1879 GCTACAATGGTGTATTTTGCACAAAGCTTAAGTAAGAAATGTTTAAACAAGGCTTAAAGT 1820
|||||
Db 181 actcaagtcacataataaaatttatcttttgcccttttacttctgaagaataatcatgctatcac 240
|||||
Cp 1819 ACTCAAGTCAATTAATAATTAATATCTTTTGGCTTTTACTTGAAGAAATCATGCTATA - 1761
|||||
Db 241 gaaatggttaagtgtcttctaataaagaagtattgtagctggaatgtgatacatgataa 300
|||||
Cp 1760 GAAATGGTTAATGTGCTTCTATAATATGAAGATATTGTAGCTGGAATGTGATACAATGTA 1701
|||||
Db 301 cagtttaagttcccatggaagtata 326
|||||
Cp 1700 CAGTTTAAGTCCCATGGAAGGTATA 1675
|||||
```

```
RESULT 15
LOCUS HUM230F10B 330 bp mRNA EST 28-AUG-1995
DEFINITION Human aorta cDNA 5'-end GEN-230F10.
ACCESSION D62000
NID 965776
KEYWORDS EST(expressed sequence tag); Human aorta; similar to none(May 29,1995).
SOURCE Homo sapiens (library: Clontech human aorta polyA+ mRNA (#6572))
CDNA to mRNA.
ORGANISM Homo sapiens
Eukaryotae; mitochondria; eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 330)
AUTHORS Fujiwara, T., Hirano, H., Katagiri, T., Kawai, A., Kuga, Y., Nagata, M.,
Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H.,
Takaichi, A., Takeda, S., Watanabe, T., Takahashi, E.-I., Hirai, Y.,
Maekawa, H., Shin, S. and Nakamura, Y.
TITLE Unpublished(303)
JOURNAL Unpublished (1995)
COMMENT Submitted (30-May-1995) to DDBJ by:
Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co., Ltd
463-10 Kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone: 0886-65-2888
Fax : 0886-37-1035.
FEATURES
source
location/Qualifiers
/organism="Homo sapiens"
/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"
BASE COUNT 116 a 57 c 32 g 118 t
ORIGIN
Query Match 15.6%; Score 316; DB 48; Length 330;
Best Local Similarity 97.8%; Pred. No. 0.00e+00;
Matches 316; Conservative 6; Mismatches 0; Indels 1; Gaps 1;
Db 9 caacaattcatcttttatcatctcaatgggaacttaaaactgttacatgtatcacattcc 68
|||||
Qy 1659 CAACAATTTCATCATTTTATACCTTCAATGGGAACCTTAAACTGTTACATGTATCACATTCC 1718
|||||
Db 69 agctacaatacttcattatttagaagcacattaaaccatttctatagcatgatttcttoa 128
|||||
Qy 1719 AGCTACAATACTTCCATTATTATAGAGCACATTAAACCAATTTCTATAGCATGATTCTTCA 1778
|||||
Db 129 agtaaaaggcaaaagataataaatttyata-ttsacttgagtagtactttaagccttggttaaa 187
|||||
Qy 1779 AGTAAAGGCAAAAGATATAAATTTTATTAATGACTTGTAGTACTTTAAGCCTTGTATAA 1838
|||||
Db . 188 acatttcttacttaacttttgcataataaaccattgttagcttacctgtaataatatacatag 247
|||||
```

```
|||||
Qy 1839 ACATTTCTTACTTAACTTTTGCRAATTTAAACCCATTGTAGCTTACCTGTAATATACATAG 1898
|||||
Db 248 tagtttaccttttaavagttgttaaaaaatattgctttaaccaaacactgttaaatatttcagat 307
|||||
Qy 1899 TAGTTTACCTTTTAAAGCTTGTAAAAATATTGCTTTAACCAACACTGTAAAAATATTTCAGAT 1958
|||||
Db 308 aaacatttatattcyygtatataa 330
|||||
Qy 1959 AAACATTATATTCTTGTATATAA 1981
|||||
```

Search completed: Tue Dec 9 09:09:41 1997
Job time : 1365 secs.

WATERFALL

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 09:10:03 1997; MasPar time 537.81 Seconds
Tabular output not generated. 1035.748 Million cell updates/sec

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCCTTCATGGCCT.....AAAAAAGGCGCCGC 2027
Comp: CTTAAGCCGAAGTACCGGA.....TTTTTTTTTTCGCGCGCG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 359085 seqs, 137405154 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS-THREE
1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204
7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210
13:EST211 14:EST212 15:EST213 16:EST214 17:EST215
18:EST216 19:EST217 20:EST218 21:EST219 22:EST220
23:EST221 24:EST222 25:EST223 26:EST224 27:EST225
28:EST226 29:EST227 30:EST228 31:EST229 32:EST230
33:EST231 34:EST232 35:EST233 36:STS1 37:STS2 38:STS3
39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10
46:STS11 47:STS12 48:STS13
EST-STS-FOUR
49:gnEST1 50:gnEST2 51:gnEST3 52:gnEST4 53:gnEST5
54:gnEST6 55:gnEST7 56:gnEST8 57:gnEST9 58:gnEST10
59:gnEST11 60:gnEST12 61:gnEST13 62:gnEST14 63:gnEST15
64:gnEST16 65:gnEST17 66:gnEST18 67:gnEST19 68:gnEST20
69:gnEST21 70:gnEST22 71:gnEST23 72:gnEST24 73:gnEST25
74:gnSTS 75:enEST1 76:enEST2 77:enEST3 78:enEST4
79:enEST5 80:enEST6 81:enEST7 82:enEST8 83:enEST9
84:enEST10 85:enEST11 86:enEST12 87:enEST13 88:enEST14
89:enEST15 90:enEST16 91:enEST17 92:enEST18 93:enEST19
94:enEST20 95:enEST21 96:enEST22 97:enEST23 98:enEST24
99:enEST25 100:enEST26 101:enEST27 102:enEST28
103:enEST29 104:enEST30 105:enEST31 106:enSTS 107:ueEST1
108:ueEST2

Database:

Statistics: Mean 12.177; Variance 2.973; scale 4.096

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description	Pred. No.
c 1	479	23.6	505	32	AA071469	ze88f06.s1 Soares fet	0.00e+00
c 2	472	23.3	505	2	W92531	ze03g10.r1 Soares fet	0.00e+00
c 3	472	23.3	505	90	HSW5314	ze03g10.r1 Soares fet	0.00e+00
c 4	366	18.1	458	33	AA105749	ml84a03.r1 Stratagene	0.00e+00
c 5	328	16.2	396	54	AA073862	mm97f07.r1 Stratagene	0.00e+00
c 6	256	12.6	262	11	AA176165	zp33h12.s1 Stratagene	0.00e+00
c 7	255	12.6	287	32	AA082155	ze88f06.r1 Soares fet	0.00e+00
c 8	255	12.6	394	101	MMAA60087	va52g08.r1 Soares mou	0.00e+00
c 9	242	11.9	523	56	AA122822	mr03c11.r1 Soares mou	0.00e+00
c 10	242	11.9	523	6	AA122822	mr03c11.r1 Soares mou	0.00e+00
c 11	209	10.3	299	94	MM3451	mb41f02.r1 Soares mou	0.00e+00
c 12	156	7.7	184	20	AA209833	mo90h01.r1 Beedington	7.57e-286
c 13	114	5.6	145	64	AA230899	mw14c07.r1 Soares mou	8.95e-198
c 14	114	5.6	145	99	MMAA30899	mw14c07.r1 Soares mou	1.03e-129
c 15	86	4.2	120	20	AA206685	zw51e09.r1 Stratagene	7.12e-86
c 16	47	2.3	502	21	AA020088	mh49d08.r1 Soares mou	2.88e-29
c 17	32	1.6	496	10	AA013810	mh06e01.r1 Soares mou	1.13e-10
c 18	31	1.5	432	29	AA030216	mh87g08.r1 Soares mou	1.47e-09
c 19	29	1.4	156	36	DM36D7S	D. melanogaster STS d	2.13e-07
c 20	28	1.4	262	11	AA176165	zp33h12.s1 Stratagene	2.36e-06
c 21	28	1.4	292	81	HS1145682	nc19a09.r1 NCI CGAP P	2.38e-06
c 22	28	1.4	292	63	AA228250	nc19a09.r1 NCI CGAP P	2.38e-06
c 23	28	1.4	376	9	AA133408	zn92f01.r1 Stratagene	2.36e-06
c 24	28	1.4	397	40	GI4542	human STS SHGC-11851	2.36e-06
c 25	28	1.4	400	40	GI4598	human STS SHGC-11851	2.36e-06
c 26	29	1.4	415	34	AA187208	zp59c07.s1 Stratagene	2.13e-07
c 27	28	1.4	505	32	AA071469	ze88f06.s1 Soares fet	2.36e-06
c 28	26	1.3	163	87	HS439337	zb96f03.s1 Soares par	2.40e-04
c 29	26	1.3	205	1	N97804	1287C3 czappfdd2.1, D	2.40e-04
c 30	26	1.3	449	100	MMAA41089	mu32h11.r1 Soares 2Nb	2.40e-04
c 31	26	1.3	449	62	AA204488	mu32h11.r1 Soares 2Nb	2.40e-04
c 32	26	1.3	449	19	AA204488	mu32h11.r1 Soares 2Nb	2.40e-04
c 33	27	1.3	482	21	AA018433	ze55g01.s1 Soares ret	2.40e-05
c 34	26	1.3	543	32	AA089167	ml85b01.r1 Stratagene	2.40e-04
c 35	26	1.3	570	101	MMAA59657	va36a10.r1 Soares mou	2.40e-04
c 36	25	1.2	158	36	DM48A3T	D. melanogaster STS d	2.19e-03
c 37	25	1.2	305	21	AA019656	ze61f11.s1 Soares ret	2.19e-03
c 38	25	1.2	327	1	N97836	2124C3 czappfdd2.1, D	2.19e-03
c 39	25	1.2	330	1	N97581	1321C3 czappfdd2.1, D	2.19e-03
c 40	25	1.2	366	67	AA238341	mx95c11.r1 Soares mou	2.19e-03
c 41	25	1.2	375	24	AA216813	mu86h01.r1 Soares mou	2.19e-03
c 42	25	1.2	379	3	AA084261	zn03e11.s1 Stratagene	2.19e-03
c 43	25	1.2	413	95	MM4855	ma53e06.r1 Soares mou	2.19e-03
c 44	25	1.2	544	13	AA186000	mc06h10.r1 Soares mou	2.19e-03
c 45	25	1.2	1171	95	MM52713	mc03e12.r1 Soares mou	2.19e-03

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESION	NID	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	AA071469	ze88f06.s1 Soares fetal heart NBHH19W Homo sapiens CDNA clone	366083 3	91578900	EST	human.	Homo sapiens	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 505)	Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.	WashU-Merck EST Project	Unpublished (1995)	Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine
	AA071469	ze88f06.s1 Soares fetal heart NBHH19W Homo sapiens CDNA clone	366083 3	91578900	EST	human.	Homo sapiens	Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 505)	Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.	WashU-Merck EST Project	Unpublished (1995)	Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:313516
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 457.

FEATURES

Location/Qualifiers

1..458

/organism="Mus musculus"

/strain="C57/B16"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTATTTT 3"

/clone="518668"

/clone_lib="Stratagene mouse kidney (#937315)"

/sex="females"

/dev_stage="4 weeks"

/lab_host="SOLR (kanamycin resistant)"

<1..458

Location/Qualifiers

119 a 140 c 116 g 83 t

mRNA

BASE COUNT

ORIGIN

Query Match 18.1%; Score 366; DB 33; Length 458;
Best Local Similarity 91.8%; Pred. No. 0.00e+00;
Matches 424; Conservative 0; Mismatches 34; Indels 4; Gaps 4;

Db 1 agcaggcggcgatgattccgctgctgaagcagtcgacccggagcaccacgaagt 60

|||||

QY 476 AGCAGCCGGCGCTTGGATCCGCTGTGATGAAGCAGTCCACCCGACACCAAGT 535

|||||

Db 61 tctgtgctgcttctgccccctgtctgtctgacgacctagatgagaccatccagcgt 120

|||||

QY 536 TCTGTGCTCGCTCTTCGCCCGCCGCTGCTGCCATGATGACCTAGACGAGACCATCCAGCAT 595

|||||

Db 121 gtacatcgctctg-gtgcaggtgaagaccgtgccccgggtcatctcgctctg-ct 178

|||||

QY 596 GCCACTCGCTCTCGGTGCAAGTGAAGACCGCTGCGCCCGGTCATGTCGCTTCGGCT 655

|||||

Db 179 tccctggccagacatgctgagtgagcagctttcccgag-acaagcactctgtcacc 237

|||||

QY 656 TCCCTGGCCGACATGCTTGAGTGGACCGTTTCCCGGACAGACGCTTTGATCC 715

|||||

Db 238 cctctgctagtag-gaccacctctgcggccacagaggagctcccaagtggtgaag 296

|||||

QY 716 CCTCTGCTAGCAGCACCACTCTCTGCCAGCCACCGAGGAAGCTTCCAAGTATGTGAAG 775

|||||

Db 297 cctgcaaaacaaagatgagacgacacacacatcatggaacccctttgtaaaatgact 356

|||||

QY 776 CTGTGCAAAATAAAATGATGATGACAACGACATAATGGAACCGCTTTGTAATAATGATT 835

|||||

Db 357 tcgactgaaatcaaagtgaagagataacgtacatcaacagagacacacaaagatcacc 416

|||||

QY 836 TTGCACGTAATAAAGTGAAGAGATAACCTATCATACACCGAGATACCAAAATCATCC 895

|||||

Db 417 tggagacaaagacagaccattacaagctgaacgctgtgt 458

|||||

QY 896 TGGAGACCAAGACAGACCATTTACAAGCTGAACGGTGTGT 937

|||||

RESULT 5 AA073862 396 bp mRNA EST 15-FEB-1997

LOCUS mm97f07.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone

DEFINITION 536389 5' similar to TR:G1151260 G1151260 TRANSMEMBRANE RECEPTOR.

; ;

ACCESSION AA073862

NID 91595602

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Mus.

REFERENCE

AUTHORS

COMMENT

TITLE

JOURNAL

JOURNAL

COMMENT

1 (bases 1 to 396)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheilenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MG1:323325

Seq primer: -28ml3 rev1 ET from Amersham

High quality sequence stop: 390.

FEATURES

source

1..396

Location/Qualifiers

/organism="Mus musculus"

/strain="NIH/Swiss"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. 93

pooled NIH/Swiss 13 day embryo hearts. Average insert

size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'

GAATTCGGCAGCAG 3' -3' adaptor sequence: 5'

CTCGAGTTTATTTT 3"

/clone="536389"

/clone_lib="Stratagene mouse heart (#937316)"

/sex="pooled"

/dev_stage="13 day embryos"

/lab_host="SOLR (kanamycin resistant)"

<1..396

Location/Qualifiers

93 a 126 c 103 g 74 t

mRNA

BASE COUNT

ORIGIN

Query Match 16.2%; Score 328; DB 54; Length 396;

Best Local Similarity 92.5%; Pred. No. 0.00e+00;

Matches 368; Conservative 0; Mismatches 28; Indels 2; Gaps 2;

Db 1 agcaggcggcgct-ggattccgctggtcatgaagcagtcgacccggagcaccacgaagt 59

|||||

QY 476 AGCAGCCGGCGCTTGGATCCGCTGTGATGAAGCAGTCCACCCGACACCAAGT 535

|||||

Db 60 tctgtgctgctcttcgccccctgtctgtctgacgacctagatgagaccatccagcgt 119

|||||

QY 536 TCTGTGCTCGCTCTTCGCCCGCCGCTGCTGCCATGATGACCTAGACGAGACCATCCAGCAT 595

|||||

Db 120 gtactgctctg-gtgcaggtgaagacgctgccccgggtcatctcgcccttcggt 178

|||||

QY 596 GCCACTCGCTTCGGTGAAGACCGCTGCGCCCGGTCATGTCGCTTCGGCT 655

|||||

Db 179 tccctggccagacatgctgagtgagcagctttcccgagacacacacacacacacac 238

|||||

QY 656 TCCCTGGCCGACATGCTTGAGTGGACCGCTTTCCCGGAGGACACGACCTTTGCAATCC 715

|||||

Db 239 cctctgctagtagcgaccacctctcgccgacacagaggagagctcccaagtggtgaag 298

|||||

QY 716 CCCTCGCTAGCAGGACCACTCTCTGCCAGCCACCGAGGAGCTCCAAGGTATGTGAAG 775

|||||

Db 299 cctgcaaaacaaagatgagacgacacacacacacacacacacacacacacacac 358

|||||

QY 776 CCTGCAAAATAAAATGATGATGACAACGACATAATGGAACCGCTTTGTAATAATGATT 835

|||||

Db 359 tcgactgaaatcaaagtgaagagataacgtacatc 396

|||||

QY 836 TTGCACGTAATAAAGTGAAGAGATAACCTATCATC 873

|||||

```

RESULT 6
LOCUS   AA176165      262 bp      mRNA      EST      30-DEC-1996
DEFINITION   zp23h12.s1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA
clone 610343 3'.
ACCESSION   AA176165
NID        91757287
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 262)
AUTHORS    Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevasakis,E.,
Waterson,R., Williamson,A., Wohlmann,P. and Wilson,R.
WashU-Merck EST Project
Unpublished (1995)

TITLE      WashU-Merck EST Project
JOURNAL
COMMENT    Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 161.

FEATURES   source
            Location/Qualifiers
            1..262
             /organism="Homo sapiens"
             /note="Vector: pBluescript SK-; Site1: EcoRI; Site2:
             XhoI; Cloned unidirectionally. Primer: Oligo dT. NT2.
             cells (Ntera-2/cl.D1) induced with Retinoic Acid for 24
             hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5'
             adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
             sequence: 5' CTCGAGTATTTTTTTTTTTTTTTT 3'"
             /clone="610343"
             /clone_lib="Stratagene neuroepithelium (#937231)"
             /dev_stage="Ntera-2/RA neuroepithelial cells"
             /lab_host="SOUR (kanamycin resistant)"
             complement(<1..>262)

BASE COUNT 100 a 27 c 40 g 91 t 4 others
ORIGIN
Query Match 12.6%; Score 256; DB 11; Length 262;
Best Local Similarity 98.1%; Pred. No. 0.00e+00;
Matches 257; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1 ggtataacaggatgaagtttatatacaagaataataatgtttatctgaaattttacag 60
Cp 2001 GGTAAACAGGATGTAAGTTTTATATACAGAAATATATGTTTATCTGAAATATTTACAG 1942

Db 61 tgttggttaaacgaatattttcacacttttaaggttaactactatgtattacaggt 120
Cp 1941 TGTGTGTTAAAGCATATTTTACACACTTTTAAAGGTAACACTACTATGTATTACAGGT 1882

Db 121 aagctacaatgggtttaatttgcacaaagttaagtaagaagaatgttttaacaaaggcttaa 180
Cp 1881 AAGCTACAATGGGTTTAATTTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCTTAAA 1822

Db 181 gtactcaagtcacataataaaattatancttngccttttacttggaagaatactctat 240
Cp 1821 GTACTCAAGTCAATATATAAAATTTATATCTTTTGCCCTTTTACTTGAAGAAATCATGCTAT 1762

Db 241 agaaatgntaatgtncctcta 262
Cp 1761 AGAAATGGTTAATGTGCTCTA 1740

RESULT 7

```

```

LOCUS   AA082155      287 bp      mRNA      EST      01-FEB-1997
DEFINITION   ze89f06.r1 Soares fetal heart NBHH19W Homo sapiens cDNA clone
366083 5'.
ACCESSION   AA082155
NID        91624408
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 287)
AUTHORS    Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevasakis,E.,
Waterson,R., Williamson,A., Wohlmann,P. and Wilson,R.
WashU-Merck EST Project
Unpublished (1995)

TITLE      WashU-Merck EST Project
JOURNAL
COMMENT    Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1171 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 262.

FEATURES   source
            Location/Qualifiers
            1..287
             /organism="Homo sapiens"
             /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
             modified polylinker; Site1: Not I; Site2: Eco RI; 1st
             strand cDNA was primed with a Not I - oligo(dT) primer [5',
             TGTTCACCACTCAAGTGGAGCGCGGCATCTTTTTTTTTTTT 3'],
             double-stranded cDNA was size selected, ligated to Eco RI
             adapters (Pharmacia), digested with Not I and cloned into
             the Not I and Eco RI sites of a modified pT7T3 vector
             (Pharmacia). Library went through one round of
             normalization to a Cot = 5. Library constructed by
             M.Fatima Bonaldo. This library was constructed from the
             same fetus as the fetal lung library, Soares fetal lung
             NBHL19W."
             /clone="366083"
             /clone_lib="Soares fetal heart NBHH19W"
             /sex="unknown"
             /dev_stage="19 weeks"
             /lab_host="DH10B (ampicillin resistant)"

BASE COUNT 62 a 84 c 85 g 51 t 5 others
ORIGIN
Query Match 12.6%; Score 255; DB 32; Length 287;
Best Local Similarity 97.5%; Pred. No. 0.00e+00;
Matches 273; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

Db 1 ggctcaagaacagcttgacgtgcacctgtgagagatgaacacatcaacgcgcctatc 60
Qy 968 GGCTCAAGACAGCTTGACGTGCACCTGTGAGGAGATGAACGACATCAACGCGCCTATC 1027

Db 61 tggctatggagacgaacagggtggggagctgtgtatcacctcgttgtaagcgggtggcaga 120
Qy 1028 TGGTCTATGGACAGAAACAGGGTGGGAGCTGGTGATCAGCTCGGTGAAGCGGTGCGAGA 1087

Db 121 aggggcagagagaggttcaacgcgcattcncggngnnnattcccaagctgagtgacttc 180
Qy 1088 AGGGGCAGAGAGAGTTCACAGCGCATCTCCCG-CAGCATCGCAAGCTGCAGTAGTATCC 1146

Db 181 cggcctcctgatggtctccacagag-cctgctccagagacagcgttgaccatttctgctcgg 239
Qy 1147 CGGCATCCTGATGGCTCCGACAGGCGCTGCTCCAGAGACGGCTGACCATTTCTGCTCGG 1206

```

QY	976	GACAGCTTGCAGTGCACCTGTGAGGAGATGAACACATCAACGCGCCCTATCTGGTCATG	1035
Db	181	ggacagaagcaggcgccgagctgggtgatcacctccgtgaaacgggtggcagaagggccag	240
QY	1036	GGACAGAACAGGGTGGGGAGCTGGTGATCACCTCGTGAAGCGGTGGCAGAGGGGCAG	1095
Db	241	agagagttcaagcgcattctcccgagcattcccgacgacgtccgaagtcgaatgtagt	289
QY	1096	AGAGAGTTCAGGCGCATCTCCCGCAGCATCCGAAGCTCGAGTGCAGT	1144
RESULT	9		
LOCUS	AA12822	523 bp	EST
DEFINITION	nr03611.r1 Soares mouse 3NbMS Mus musculus cDNA clone 596372 5'		16-FEB-1997
ACCESSION	AA12822		
NID	gi681779		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
	Eukaryotae; mitochondrional eukaryotes; Metazoa; Chordata;		
	Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Mus.		
REFERENCE	1 (bases 1 to 523)		
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.		
TITLE	The WashU-HMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicinep 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:361804 Seq primer: -28M13 rev2 from Amersham High quality sequence stop: 469.		
FEATURES	Location/Qualifiers		
SOURCE	1..523 /organism="Mus musculus" /strain="C57BL/6J" /note="Vector: pT7R3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCGAAGTGGGAGCGCGCGCTGTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through three rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo." /clone="596372" /clone_lib="Soares mouse 3NbMS" /sex="male" /tissue_type="Spleen" /dev_stage="4 weeks" /lab_host="DH10B" <1..523		
BASE COUNT	126 a 142 c 129 g 126 t		
ORIGIN			
Query Match	11.9%;	Score 242;	DB 56; Length 523;
Best Local Similarity	80.28;	Pred.No. 0.00e+00;	
Matches	408;	Conservative 0;	Mismatches 88; Indels 13; Gaps 5;
Db	1	tcogtgcgtgctcaagacagcctgcagtcgacctgtgagagatgaacgacatcaac	60

QY 958 TCGGTGCTGTGGCTCAAGACAGCTTGCAGTGCACCTGTGTGAGGAGATGAACGACATCAAC 1017
 Db 61 gctcgtatcttggttcattggtgacagagagggcgagcctggtgtagctcactcgtgaa 120
 QY 1018 GCGCCCTATCTGCTCATGGGACAGAAACAGGGTGGGAGC-TGGTGATCACCTCGGTGAA 1076
 Db 121 acgtgtggcagagggccagagaggttcaagcgcatctcccgacacatccgcaactgca 180
 QY 1077 GCGGTGGCAGAGGGGACAGAGAGTTCAGGCGCATCTCCCGCAGCATCCGCAAGCTGCA 1136
 Db 181 atgtagtttccagtggtggtctctctccatccagggccctgagctgtgtagaccact 240
 QY 1137 GTGCTAGTCCCGGATCTGATGGCTCCGACAGAGGCTGCTCCAGACGAGGGTGACCAT 1196
 Db 241 tccgtccgcgacctatttccgggttcccaagcacagtcctccgaaagctacagccccag 300
 QY 1197 TCTGCTCCGGATCTCAGCTCCGTTCCCAAGCACACTCCTAGCT-GCTCCAGTCTCAG 1255
 Db 301 cttggagcgttgcctcgtcctcgtgtagtctccatccctaacatgctcctgagttataa 360
 QY 1256 CTTGGGACAGTCCCGCTTTCACGTTTGCAGTTTGCATCCCCAGCATTTCTGAGTTATA 1315
 Db 361 ggccttagggccttgaaacccatagctgtttccaggaagcgaaagccccatccag 420
 QY 1316 GCCACAGGAG----TGA-----TAGCTGTTTCACTAAAG-GAAAGGCCACCCGA 1364
 Db 421 atctgtacaaatattcaaaactataaaatcatgactattttatgaagtttttagaacag 480
 QY 1365 ATCTGTAGAAATATTCAAACTAATAAATCATGAATATTTTATGAGTTTAAAAATAG 1424
 Db 481 ctgctttaaggttagttttgaatagctg 509
 QY 1425 CTCACTTTAAAGCTAGTTTTCGAATAGGTG 1453

RESULT 10
 LOCUS AAL22822 523 bp mRNA EST 21-NOV-1996
 DEFINITION mr03c11.r1 Soares mouse 3NDMS Mus musculus cDNA clone 596372 5'
 ACCESSION AAL22822
 NID g1681779
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Mus.
 REFERENCE 1 (bases 1 to 523)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and
 Waterston,R.
 TITLE The WashU-HHMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:361804
 Seq primer: -28M13 rev2 from AmerSham
 High quality sequence stop: 469.
 Location/Qualifiers
 1..523
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; [5'

TGTTACCAATCTGAAGTGGAGCGCGCGCTGCTTTTTTTTTTTTTTTTTTTTTTTTTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified p7T3 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through
 three rounds of normalization, and was constructed by
 Bento Soares and M.Fatima Bonaldo."
 /clone="596372"
 /clone_lib="Soares mouse 3NDMS"
 /sex="male"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 BASE COUNT 126 a 142 c 129 g 126 t
 ORIGIN
 Query Match 11.9%; Score 242; DB 6; Length 523;
 Best Local Similarity 80.2%; Pred.No. 0.00e+00;
 Matches 408; Conservative 0; Mismatches 88; Indels 13; Gaps 5;
 Db 1 tccgtgctgtggtctcaaaagacagcctgcagtcacctgtgagagatgaacgacatcaac 60
 QY 958 TCGGTGCTGTGGCTCAAGACAGCTTGCAGTGCACCTGTGTGAGGAGATGAACGACATCAAC 1017
 Db 61 gctcgtatcttggttcattggtgacagagggcgagcctggtgtagctcactcgtgaa 120
 QY 1018 GCGCCCTATCTGCTCATGGGACAGAAACAGGGTGGGAGC-TGGTGATCACCTCGGTGAA 1076
 Db 121 aggtgtggcagagggccagagagaggttcaagcgcatctcccgacacatcccgagctgca 180
 QY 1077 GCGGTGGCAGAGGGGACAGAGAGTTCAGGCGCATCTCCCGCAGCATCCGCAAGCTGCA 1136
 Db 181 atgtagtttccagtggtggtggtctctccatccagggccctgagctgtgtagaccact 240
 QY 1137 GTGCTAGTCCCGGATCTGATGGCTCCGACAGGCTGCTCCAGACGAGGGTGACCAT 1196
 Db 241 tccgtccgcgacctatttccgggttcccaagcacagtcctccgaaagctacagccccag 300
 QY 1197 TCTGCTCCGGATCTCAGCTCCGTTCCCAAGCACACTCCTAGCT-GCTCCAGTCTCAG 1255
 Db 301 cttggagcgttgcctcgtcctcgtgtagtctccatccctaacatgctcctgagttataa 360
 QY 1256 CTTGGGACAGTCCCGCTTTCACGTTTGCAGTTTGCATCCCCAGCATTTCTGAGTTATA 1315
 Db 361 ggccttagggccttgaaacccatagctgtttccaggaagcgaaagccccatccag 420
 QY 1316 GCCACAGGAG----TGA-----TAGCTGTTTCACTAAAG-GAAAGGCCACCCGA 1364
 Db 421 atctgtacaaatattcaaaactataaaatcatgactattttatgaagtttttagaacag 480
 QY 1365 ATCTGTAGAAATATTCAAACTAATAAATCATGAATATTTTATGAGTTTAAAAATAG 1424
 Db 481 ctgctttaaggttagttttgaatagctg 509
 QY 1425 CTCACTTTAAAGCTAGTTTTCGAATAGGTG 1453
 RESULT 11
 ID MM3451 standard; RNA; EST; 299 BP.
 AC W08345;
 NI g1282366
 DT 27-APR-1996 (Rel. 47, Created)
 DE 05-MAR-1997 (Rel. 51, Last updated, Version 2)
 DE mb41f02.r1 Soares mouse p3NM19.5 Mus musculus cDNA clone 331995
 DE 5'.
 KW EST.
 OS Mus musculus (house mouse)
 OC Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;
 OC Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 OC Mus.
 RN [1]
 RP 1-299
 RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,

RA	Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA	Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA	Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA	Waterston R.;
RT	"The WashU-HHMI Mouse EST Project";
RL	Unpublished.
CC	Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC	Washington University School of Medicinep 4444 Forest Park Parkway,
CC	Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC	Email: mouseest@watson.wustl.edu This clone is available
CC	royalty-free through LNL ; contact the IMAGE Consortium
CC	(info@image.lnl.gov) for further information. MGI:213395 Seq
CC	primer: mob.REGA+ET High quality sequence stop: 280.
CC	Key
FH	Location/Qualifiers
FT	1..299
FT	/organism="Mus musculus"
FT	/note="Vector: pT7T3D (Pharmacia) with a modified
FT	polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
FT	was primed with a Not I - oligo(dT) primer [5',
FT	TGTTACCATCTGAAGTGGAGCGCGGCAATTTTCTTTTCTTTT 3'],
FT	double-stranded cDNA was size selected, ligated to Eco RI
FT	adapters (Pharmacia), digested with Not I and cloned into
FT	the Not I and Eco RI sites of a modified pT7T3 vector
FT	(Pharmacia). Library went through one round of
FT	normalization to a Cot = 5. Library constructed by Bento
FT	Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr
FT	Minoru Ko (Wayne State University)."
FT	/clone="331995"
FT	/clone_lib="Scores mouse p3NNMF19.5"
FT	/dev_stage="19.5 dpc total fetus"
FT	/lab_host="DHI0B (ampicillin resistant)"
FT	<1..>299
FT	mRNA
FT	Sequence 299 BP: 79 A; 95 C; 72 G; 53 T; 0 other;
FT	

Query Match	10.3%	Score 209;	DB 94;	Length 299;
Best Local Similarity	89.8%;	Pred. No. 7.57e-286;		
Matches	264;	Conservative 0;	Mismatches 25;	Indels 5; Gaps 5;
Db	11	gcctgtcaactcgtctt-cgtgcag-tgaaggaccctgcgcgc-ggtcatgtccgctt	67	
QY	591	GCCATGCCACTCCTCTGCTGAGGTGAAGGACCGCTGCGCCCGGTGATGTCGCCCTT	650	
Db	68	cg-cttccccctggccagacatcgtgaagtgcgaccctttccgcag-acaaagacacctg	125	
QY	651	CGGCTTCCCTTGCCCCGACATGCTTGAGTCGACCGCTTCCCCCGACGAACGACCTTG	710	
Db	126	catccccctcgtctagtagcaccacctcctgcgcgcgcacagaggaagctcccaaggtg	185	
QY	711	CATCCCCCTCGCTAGCAGCACCACCTCTCGCAGCCACCGAGGAGCTCCAAAGGTATG	770	
Db	186	tgaagctgcacaaacccaagaatgagcagacacacatcatcgtgaacacctttgtataaa	245	
QY	771	TGAAGCGCTGCAAAATAAAAAATGATGATGACACGACATAATGGAACGCTTTGTAAAAA	830	
Db	246	tgacttcgcactgaataatcaaaagtgaaggagataacggtacatcaacacagagacac	299	
QY	831	TGATTTTGCACTGAAAATAAAAGTGAAGGAGATTAACCTACATCAACCGGAGATC	884	

RESULT	12	AA209833	184 bp	mRNA	EST	29-JAN-1997
LOCUS		mo90h01.r1	Beddington mouse embryonic region			
DEFINITION		clone 560831 5'				
ACCESSION		AA209833				
NID		g1808152				
KEYWORDS		EST.				
SOURCE		house mouse.				
ORGANISM		Mus musculus				
		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;				
		Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
		Mus.				

REFERENCE AUTHORS	TITLE JOURNAL COMMENT
1. [Empty]	[Empty]
2. [Empty]	[Empty]
3. [Empty]	[Empty]
4. [Empty]	[Empty]
5. [Empty]	[Empty]
6. [Empty]	[Empty]
7. [Empty]	[Empty]
8. [Empty]	[Empty]
9. [Empty]	[Empty]
10. [Empty]	[Empty]
11. [Empty]	[Empty]
12. [Empty]	[Empty]
13. [Empty]	[Empty]
14. [Empty]	[Empty]
15. [Empty]	[Empty]
16. [Empty]	[Empty]
17. [Empty]	[Empty]
18. [Empty]	[Empty]
19. [Empty]	[Empty]
20. [Empty]	[Empty]
21. [Empty]	[Empty]
22. [Empty]	[Empty]
23. [Empty]	[Empty]
24. [Empty]	[Empty]
25. [Empty]	[Empty]
26. [Empty]	[Empty]
27. [Empty]	[Empty]
28. [Empty]	[Empty]
29. [Empty]	[Empty]
30. [Empty]	[Empty]
31. [Empty]	[Empty]
32. [Empty]	[Empty]
33. [Empty]	[Empty]
34. [Empty]	[Empty]
35. [Empty]	[Empty]
36. [Empty]	[Empty]
37. [Empty]	[Empty]
38. [Empty]	[Empty]
39. [Empty]	[Empty]
40. [Empty]	[Empty]
41. [Empty]	[Empty]
42. [Empty]	[Empty]
43. [Empty]	[Empty]
44. [Empty]	[Empty]
45. [Empty]	[Empty]
46. [Empty]	[Empty]
47. [Empty]	[Empty]
48. [Empty]	[Empty]
49. [Empty]	[Empty]
50. [Empty]	[Empty]
51. [Empty]	[Empty]
52. [Empty]	[Empty]
53. [Empty]	[Empty]
54. [Empty]	[Empty]
55. [Empty]	[Empty]
56. [Empty]	[Empty]
57. [Empty]	[Empty]
58. [Empty]	[Empty]
59. [Empty]	[Empty]
60. [Empty]	[Empty]
61. [Empty]	[Empty]
62. [Empty]	[Empty]
63. [Empty]	[Empty]
64. [Empty]	[Empty]
65. [Empty]	[Empty]
66. [Empty]	[Empty]
67. [Empty]	[Empty]
68. [Empty]	[Empty]
69. [Empty]	[Empty]
70. [Empty]	[Empty]
71. [Empty]	[Empty]
72. [Empty]	[Empty]
73. [Empty]	[Empty]
74. [Empty]	[Empty]
75. [Empty]	[Empty]
76. [Empty]	[Empty]
77. [Empty]	[Empty]
78. [Empty]	[Empty]
79. [Empty]	[Empty]
80. [Empty]	[Empty]
81. [Empty]	[Empty]
82. [Empty]	[Empty]
83. [Empty]	[Empty]
84. [Empty]	[Empty]
85. [Empty]	[Empty]
86. [Empty]	[Empty]
87. [Empty]	[Empty]
88. [Empty]	[Empty]
89. [Empty]	[Empty]
90. [Empty]	[Empty]
91. [Empty]	[Empty]
92. [Empty]	[Empty]
93. [Empty]	[Empty]
94. [Empty]	[Empty]
95. [Empty]	[Empty]
96. [Empty]	[Empty]
97. [Empty]	[Empty]
98. [Empty]	[Empty]
99. [Empty]	[Empty]
100. [Empty]	[Empty]

1 (bases 1 to 184)
Marra,N., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le.M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1965, Vol. 68, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Technology	1978, Vol. 1, No. 2, pp. 1-10
3. The Importance of Parental Involvement	Journal of Educational Psychology	1985, Vol. 77, No. 3, pp. 1-10
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1990, Vol. 93, No. 1, pp. 1-10
5. The Role of the Teacher in the Classroom	Journal of Educational Research	1995, Vol. 98, No. 1, pp. 1-10
6. The Impact of Technology on Education	Journal of Educational Technology	2000, Vol. 3, No. 2, pp. 1-10
7. The Importance of Parental Involvement	Journal of Educational Psychology	2005, Vol. 97, No. 3, pp. 1-10
8. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	2010, Vol. 113, No. 1, pp. 1-10
9. The Role of the Teacher in the Classroom	Journal of Educational Research	2015, Vol. 118, No. 1, pp. 1-10
10. The Impact of Technology on Education	Journal of Educational Technology	2020, Vol. 6, No. 2, pp. 1-10

Contact: Marra M/Mouse EST Project
WASHU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Fax: 314 200 1010
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the
 email: mouseestewatson.wustli.edu
 IMAGE Consortium (info@image.llnl.gov) for further information.

MG1:341673
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 147.

FEATURES	SOURCE
1. Feature 1	Source 1
2. Feature 2	Source 2
3. Feature 3	Source 3
4. Feature 4	Source 4
5. Feature 5	Source 5
6. Feature 6	Source 6
7. Feature 7	Source 7
8. Feature 8	Source 8
9. Feature 9	Source 9
10. Feature 10	Source 10

```

1.200
/organism="Mus musculus"
/strain="C57BL6 x DBA"
/note="vector: pCMV-SPORT; Site_1: SalI; Site_
2 cloned unidirectionally. Primer: oligo dT. GATC
3 embryos were collected at 7.5dpc from C57BL6 x
4 matings, excluding embryos that had developed
5 and all extraembryonic tissues. Average insert
6 kb (range: 0.5 - 3.0 kb). Referenced in Develop
7 2479-2489 (1995)."
/clone="560881"
/clone_lib="beddington mouse embryonic region"
/sex="pooled"
/dev_stage="7.5dpc"
/lab_host="DH12S"

```

BASE COUNT	61 a	45 c	48 g	30 t
ORIGIN				
mpRNA				
rad_posc-Dr125				
		<1..>184		

Query Match 7.7%; Score 156; DB 20; Length 184;
Best Local Similarity 94.0%; Pred. No. 8.95e-198;
Matches 172. Conservative 0. Mismatches 10. Indels

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

001 gataacgtacatcaacagacacacccaagatcctctggagacaaagagcaagaccattta 60

861 GATAACCTACATCAACCGAGATACCAAAATCATCTGGAGACCAAGAGCAAGACCATTA 92

61 caagctgaacggcggtgcgaaagg-acctgaagaaatcgtgctgtggctcaaaagacag 11

921 CAAGCTGAACGGTGTGTCCGAAGGGACCTGAAGAAATCGGTGCTGTGGCTCAAAGACAG 98

120 cctgcagtgacctgtgagagatgaacagacatcaacgctccgatatctggtcatgggaca

981 CTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACGGGCCCTATCTGGTCATGGGACA 10

180 gaa 182

1041 GAA 1043

RESULTS

RESULTS	13
LOCUS	AA230899 145 bp mRNA EST
DEFINITION	mwl4c07.r1 Soares mouse 3NME12 5' Mus musculus cdna clone 670668 5'
ACCESSION	AA230899
NID	g1853194
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
Eukaryotae;	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata;	Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.	Mus.

REFERENCE
AUTHORS

1 (bases 1 to 145)
Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T.,
Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,
Schellenberg K., Steptoe M., Tan F., Underwood K., Moore B.,
Theising B., Wylie T., Lennon G., Soares B., Wilson R. and
Waterston R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)

TITLE
JOURNAL
COMMENT

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:410372
Putative full length read
vector to vector length is 146
Seq primer: -28m13 rev2 ET from Amersham.

FEATURES

source

1..145
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGGCGCGCTTATTTTATTTT 3'],
on total mouse RNA [provided by Minoru Ko, Wayne State
Univ.]; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/clone="670668"
/clone_lib="Soares mouse 3NME12 5"
/sex="unknown"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
1..145
51 a 22 c 19 g 53 t

BASE COUNT
ORIGIN

Query Match 5.6%; Score 114; DB 64; Length 145;
Best Local Similarity 92.5%; Pred. No. 1.03e-129;
Matches 135; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
Db 1 gcaaaactaaatc-attgtagcttacctgtaataacgtagtagtgacctggaaaagtgg 59
Qy 1859 GCAAAATTAACCCATGTAGCTGTAATATACATAGTAGTTTACCTTTAAAGTTG 1918
Db 60 taaaaatattgcttttaaccgacactgtaataatttcagataaacattattcttggat 119
Qy 1919 TAAAAATATTGCTTTTAAACCAACACTGTAATATTTCAGATAAACATTATTTCTT-GTAT 1977
Db 120 ataaactttacatctgttgtaacctta 145
Qy 1978 ATAACTTTACATCTGTTTACCTTA 2003

RESULT 14

ID MAA30899 standard; RNA; EST; 145 BP.
AC AA230899;
NI 91853194
DT 27-FEB-1997 (Rel. 51, Created)
DT 27-FEB-1997 (Rel. 51, Last updated, Version 1)
DE mw14c07.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 670668 5'.
KW EST.
OS Mus musculus (house mouse)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

OC Mus.

RP 1-145
RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
Waterston R.;
RA "The WashU-HHMI Mouse EST Project";
RT Unpublished.
RL Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC Washington University School of Medicine 444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: mouseest@watson.wustl.edu This clone is available
CC royalty-free through LNL; contact the IMAGE Consortium
CC (info@image.llnl.gov) for further information. MGI:410372 Putative
CC full length read vector to vector length is 146 Seq primer: -28m13
CC rev2 ET from Amersham.
CF Key Location/Qualifiers
FH 1..145
FT source
FT /organism="Mus musculus"
FT /strain="C57BL/6J"
FT /note="Vector: pT73D-Pac (Pharmacia) with a modified
FT polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT was primed with a Not I - oligo(dT) primer [5',
FT TGTTACCAATCTGAAGTGGGCGCGCTTATTTTATTTT 3'], O
FT total mouse RNA [provided by Minoru Ko, Wayne State Univ.];
FT double-stranded cDNA was ligated to Eco RI adaptors
FT (Pharmacia), digested with Not I and cloned into the Not I
FT and Eco RI sites of the modified pT73 vector. Library went
FT through one round of normalization, and was constructed by
FT Bento Soares and M. Fatima Bonaldo."
FT /clone="670668"
FT /clone_lib="Soares mouse 3NME12 5"
FT /sex="unknown"
FT /dev_stage="12.5dpc total fetus"
FT /lab_host="DH10B"
FT 1..145
FT mRNA
FT Sequence 145 BP; 51 A; 22 C; 19 G; 53 T; 0 other;
SQ
Query Match 5.6%; Score 114; DB 99; Length 145;
Best Local Similarity 92.5%; Pred. No. 1.03e-129;
Matches 135; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
Db 1 gcaaaactaaatc-attgtagcttacctgtaataacgtagtagtgacctggaaaagtgg 59
Qy 1859 GCAAAATTAACCCATGTAGCTGTAATATACATAGTAGTTTACCTTTAAAGTTG 1918
Db 60 taaaaatattgcttttaaccgacactgtaataatttcagataaacattattcttggat 119
Qy 1919 TAAAAATATTGCTTTTAAACCAACACTGTAATATTTCAGATAAACATTATTTCTT-GTAT 1977
Db 120 ataaactttacatctgttgtaacctta 145
Qy 1978 ATAACTTTACATCTGTTTACCTTA 2003
RESULT 15
LOCUS AA206685 120 bp mRNA EST 27-JAN-1997
DEFINITION Zg51e09.r1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA
clone 645160 5'.
ACCESSION AA206685
NID g1802055
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 120)
AUTHORS Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E.,
Waterson, R., Williamson, A., Wohlmann, P. and Wilson, R.
WashU-Merck EST Project
Unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilton RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 67.

FEATURES

SOURCE

1..120
/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. NT2
cells (Ntera-2/cl.D1) induced with Retinoic Acid for 24
hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGCGCAGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"
/clone="645160"
/dev_stage="Ntera-2/RA neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
<1..>120

BASE COUNT
ORIGIN

12 a 54 c 35 g 18 t 1 others

Query Match 4.28; Score 86; DB 20; Length 120;
Best Local Similarity 89.38; Pred. No. 7.12e-86;
Matches 108; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

Db 2 cccccggagctg-cg-gcggttgagcncctgcgcgcctgtctcccggtgtccgct 60

QY 48 CCCCCGGAGCTCGCGGGGCTTGCAGCGCTCGCCCGGCTCCTCCCGGTGTCGGCT 107

Db 61 tctccgccccagccgcgg-tgccagcttttcggggccccagtcgcaaccaacgaag 119

QY 108 TCTCCGCGCCCAAGCGCGCGGCTGCCAGCTTTTCGGGGCCCGAGTCGCACCCAGCGAAG 167

Db 120 a 120

QY 168 A 168

Search completed: Tue Dec 9 09:22:41 1997.
Job time : 758 secs.

QY	521	CGGACACCAAGAGTTCCTGCTGCTCTTCCGCCCCCTCTGCTCGATGACCTAGAC	580
Db	619	agccgtgcccctgcgcctcgtggtgtagagagcgccaaagccggctgctgcgcgctga	678
QY	581	AGACCATCCAGCCATGCCACTCGCTCTGCTGTCAGGTGAAGACCCGCTGCCCCGGTCA	640
Db	679	tgcgcagtaacggcttcgcctgtggccgagcgatgagctgcagccgctctccc	730
QY	641	TGNCGCGCTTCGGCTTCCTCGCCGCGACATGCTTCAGTGGCAGCCGTTTCCC	692

3	MMU43321	2421 bp	DNA	ROD	24-FEB-1996
LOCUS	Mus musculus putative transmembrane receptor (frizzled 8) gene, complete cds.				
DEFINITION					
ACCESSION	U43321				
NID	91151259				
KEYWORDS	house mouse.				
SOURCE	Mus musculus				
ORGANISM	Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 2421)				
AUTHORS	Wang, Y., Macke, J.P., Abella, B.S., Andreasson, K., Worley, P., Gilbert, D.J., Copeland, N.G., Jenkins, N.A. and Nathans, J.				
TITLE	A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled				
JOURNAL	J. Biol. Chem. 271 (8), 4468-4476 (1996)				
MEDLINE	96224032				
REFERENCE	2 (bases 1 to 2421)				
AUTHORS	Abella, B., Wang, Y., Macke, J.P. and Nathans, J.				
TITLE	Direct Submission				
JOURNAL	Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205				

Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205

Location/Qualifiers

1. 2421

/organism="Mus musculus"

/chromosome="18"

/map="between Tpl2 and Cdh2"

188.2245

/gene="frizzled 8"

/note="putative transmembrane receptor"

/codon_start=1

/product="transmembrane receptor"

/db_xref="PID:gl151260"

/translation="MEWGVLLEVTSLAALAVLQRSGGAAASAKELACQETVPLCK
GIGYNTYNNFNDTODEAGLEHQFWPLVEIQSPDLKFFLCSTMTPICLEDTYK
PLPPCRSEKARAKAGPLMROYGFAMPDMRCDRLEPQGNPDTLICMDYRDLTAA
PSPPRPLPPPPQGPQSPGSGHSPRPPGARPPIRGSGSRGSDAAAPSPRGKARP
GGGAPECGCQCAPMWSVSSEHPLYNRYTQIANCALPCHNPPFFSQDERAFTVF
WGLMSVLGFSTFTFTVFLYIDMERFKYPERPIIFLSACIYLFVSGVILRVLVAGHEK
VACSGAPGAGRGAGGAGAAAGAGAGSGAGSGARGATELGAQEVRVETTGPA
LCTVFLVLYFPGMASSIIWVILSLTWELAAAGMGCNEAIGYSQTFHLAAVLVPSVK
STAVLASVSDGDPVAGICVYGNQSLDNLRGVFLAPLYVILIGTFLAGFVSLERI
RSVYKQGGGTHDKLEKMLRGLGFLTITVPAAVVACLFEQHNRRPWEATHWCKP
LUDLQPDQARPDYAVENMLKTFMCLVVGITSGVWVWSGTLTRCALTRCCWAKSG
AAVGAGAGSGPGGPGGGGGHGGGGGSLYSDVYSTGLTWRSGTASSVSYPKQMLP
SQV"

403 a 772 c 779 g 467 t

BASE COUNT
ORIGIN

	Query Match	2.88;	Score 56;	DB 87;	Length 2421;
	Best Local Similarity 59.4%;				
	Matches 177;	Conservative	Pred. No. 1.17e-17;	Mismatches 121;	Indels 0; Gaps
Db	306	tgcggttgcaaaagcatcggttacaa	tacactatadagcccaaccagttcaaacag	365	
QY	395	TCAGGTGTGCCACGCATCGATACAG	AACATGGCGTGTGCCCAACCTGCTGGGCCACG	454	
Db	366	abacgcaagaatagggcggggtctaga	ggtgcaccagttttggccgctgggtgggagatacagt	425	

it 11 10 11

.

FC

•

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPerch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Dec 9 08:30:17 1997; MacPar time 741.29 Seconds
Tabular output not generated.
1311.186 Million cell updates/sec

Title: >848439-2-trans.rge
Description: (1-885) from translate.seq
Perfect Score: 734
N.A. Sequence: 1 ATGTCNARGCNGCNGNS.....SNATNMGNAARYTNCARTGY 885
Comp: TACRANGTYCNGCNGCNSW.....WNTANKNTTYRANGTYACR

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPerch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 08:30:17 1997; MacPar time 741.29 Seconds
Tabular output not generated.
1311.186 Million cell updates/sec

Tabular output not generated.

Title: >848439-2-trans.rge

Description: (1-885) from translate.seq

Perfect Score: 734

N.A. Sequence: 1 ATGTCNARGCNGCNGNS.....SNATNMGNAARYTNCARTGY 885

Comp: TACRANGTYCNGCNGCNSW.....WNTANKNTTYRANGTYACR

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 362067 seqs, 549138275 bases x 2

Post-processing: Minimum Match 04

Listing first 45 summaries

Database: embl-new3

1-BCT 2-FUN 3-GEN1 4-GEN2 5-HTG1 6-HTG2 7-HUM 8-INV

9-ORG 10-MAM 11-VRT 12-PLN 13-PRO 14-ROD 15-SYN 16-UNC

17-VIR

genbank99

18-BCT1 19-BCT2 20-BCT3 21-BCT4 22-BCT5 23-BCT6 24-BCT7

25-BCT8 26-BCT9 27-BCT10 28-BCT11 29-GEN1 30-GEN2

31-GEN3 32-HTG1 33-HTG2 34-HTG3 35-INV1 36-INV2 37-INV3

38-INV4 39-INV5 40-INV6 41-INV7 42-INV3 43-INV9 44-INV10

45-INV11 46-MAM1 47-MAM2 48-MAM3 49-VRT1 50-VRT2 51-VRT3

52-VRT4 53-PAT1 54-PAT2 55-PAT3 56-PAT4 57-PAT5 58-PHG

59-PLN1 60-PLN2 61-PLN3 62-PLN4 63-PLN5 64-PLN6 65-PLN7

66-PLN8 67-PLN9 68-PLN10 69-PLN11 70-PR11 71-PR12

72-PR13 73-PR14 74-PR15 75-PR16 76-PR17 77-PR18 78-PR19

79-PR10 80-PR11 81-PR12 82-PR13 83-PR14 84-PR15

85-ROD1 86-ROD2 87-ROD3 88-ROD4 89-ROD5 90-ROD6 91-ROD7

92-ROD8 93-STR 94-SYN 95-UNA 96-VRL1 97-VRL2 98-VRL3

99-VRL4 100-VRL5 101-VRL6 102-VRL7 103-VRL8 104-VRL9

105-VRL10

genbank-new3

106-BCT 107-GEN1 108-GEN2 109-HTG1 110-HTG2 111-INV

112-MAM 113-VRT 114-PHG 115-PLN 116-PR11 117-PR12

118-ROD 119-SYN 120-UNA 121-VRL

Database: u-embl50_99

122-part1

Statistics: Mean 15.331; Variance 11.948; scale 1.283

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	720	98.1	1776 85	D50462	Mouse SDF5 mRNA, comp	0.00e+00
2	89	12.1	2421 87	MMU43321	Mus musculus putative	1.25e-22
3	75	10.2	2334 77	HSU43318	Human putative trans	3.36e-16
4	64	8.7	1770 38	CEU43316	Caenorhabditis elegans	2.45e-11
5	63	8.6	2831 87	MMU43317	Mus musculus putative	6.61e-11
6	31	8.3	2085 40	DMU65589	Drosophila melanogaster	4.75e-10
7	59	8.0	2260 87	MMU43320	Mus musculus putative	3.36e-09
8	56	7.6	1913 38	CEU63557	Caenorhabditis elegans	6.05e-08
9	56	7.6	4540 90	RATFRZ2H	Rattus norvegicus Dro	6.05e-08
10	54	7.4	1923 80	HUMFRIZ	Human frizzled gene p	4.04e-07
11	52	7.1	1912 90	RATFRZ2H	Rattus norvegicus Dro	2.64e-06
12	52	7.1	34576 8	CET2308	Caenorhabditis elegans	2.64e-06
13	52	7.1	34576 38	CET2308	Caenorhabditis elegans	2.64e-06
14	52	7.1	34576 111	CET2308	Caenorhabditis elegans	2.64e-06
15	49	6.7	215 57	I28278	Sequence 5 from paten	4.18e-05
16	47	6.4	215 57	I28278	Sequence 5 from paten	2.54e-04
17	46	6.3	2624 87	MMU43205	Mus musculus frizzled	6.19e-04
18	44	6.0	354 8	OFU89259	Oxytricha fallax 57kd	3.59e-03
19	44	6.0	354 111	OFU89259	Oxytricha fallax 57kd	3.59e-03
20	43	5.9	354 111	OFU89259	Oxytricha fallax 57kd	8.51e-03
21	43	5.9	354 8	OFU89259	Oxytricha fallax 57kd	8.51e-03
22	39	5.3	74911 33	HS209H1	Human DNA sequence **	2.45e-01
23	38	5.2	370 8	OFU89262	Oxytricha fallax 57kd	5.52e-01
24	38	5.2	370 111	OFU89262	Oxytricha fallax 57kd	5.52e-01
25	37	5.0	1509 23	INSREP	Insertion sequence IS	1.23e+00
26	37	5.0	1513 24	MU01217	Mycoplasma hyorhinis	1.23e+00
27	37	5.0	1518 23	INSGLA	Insertion sequence IS	1.23e+00
28	37	5.0	2187 43	PEAHSF70H	Plasmodium falciparum	1.23e+00
29	37	5.0	30864 38	CET01H10	Caenorhabditis elegans	1.23e+00
30	36	4.9	2492 87	MMU43319	Mus musculus putative	2.70e+00
31	36	4.9	140571 34	HS46H23	Human DNA sequence **	2.70e+00
32	35	4.8	1291 11	XLU78598	Xenopus laevis Frzb p	5.86e+00
33	35	4.8	1291 113	XLU78598	Xenopus laevis Frzb p	5.86e+00
34	35	4.8	1975 52	XLU68059	Xenopus laevis frezli	5.86e+00
35	35	4.8	4284 24	MHV143C	M. hyorhinis vlpA, vlp	5.86e+00
36	35	4.8	3200 36	CEF54B3	Caenorhabditis elegans	5.86e+00
37	35	4.8	47745 33	CET26E3	*** SEQUENCING IN PRO	5.86e+00
38	34	4.6	294 83	HUMUT612B	Human STS UT612, 3' p	1.25e+01
39	34	4.6	413 44	PFU17812	Periplaneta fuliginos	1.25e+01
40	34	4.6	1054 39	DDU67923	Dictyostellium discoid	1.25e+01
41	34	4.6	3618 38	CEZYG11	Caenorhabditis elegans	1.25e+01
42	34	4.6	4413 69	YSPCRM1N1	Yeast cml-N1 gene fo	1.25e+01
43	34	4.6	13815 88	MUSDYSA	Mouse dystrophin mRNA	1.25e+01
44	34	4.6	34731 78	HSU75931	Human Xp22 cosmid U25	1.25e+01
45	34	4.6	119704 63	OSCHLP1XX	O. sinensis chloroplas	1.25e+01

ALIGNMENTS

RESULT 1 D50462 1776 bp mRNA 19-DEC-1996
LOCUS Mouse SDF5 mRNA, complete cds.
DEFINITION

TITLE A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled

JOURNAL J. Biol. Chem. 271 (8), 4468-4476 (1996)

MEDLINE 96224032

REFERENCE 2 (bases 1 to 2421)

AUTHORS Abella,B., Wang,Y., Macke,J.P. and Nathans,J.

TITLE Direct Submission

JOURNAL Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205

FEATURES

source

1..2421

/organism="Mus musculus"

/chromosome="18"

/map="between Tpl2 and Cdh2"

188..2245

/genes="frizzled 8"

/notes="putative transmembrane receptor"

/codon_start=1

/product="transmembrane receptor"

/db_xref="PID:g1151260"

/translation="MFWGYLEVTSLLAALAVLQSSGAASAKELACQELTPVLPCK LGNYTYMPNQNHDTQDEAGLEVHQFWPLVEIQSPDLKFFLCISWYTPICLEDYK PLPPCSVCERAKACGSLMRQYGFAPWRMCRDLPEQGNPDILCMDYNRDLITAA PSPRLRPPPPGQPPSGCHSRPPGARPHRGSGSGCSDAAAPPSRGKARPP GGAAFCFPGQCRADPMVSVSSERHPLYNRVKTGTANCALEPCNPFSSQDERAFVFWIGLWSLFCFYSTFATVSTFLIDMERFYPERPIIFLSACYLFVSUGYLVRLVAGHEK VACSGNAGCGCGAGGAGGAAGAGAGGASSPGARGEYELGAVEQHVRYETGPA LCTWELVLYFFGNASSLHWLISLTFWLAAGMGNENAIAGTSQYFHLAHLVPSVK SIAYLALSDVDPVAGICTYGNQSDLNIRGLFVLAPLVLYLFTGMEFLAGVSLRI RSVIKQGGPRTKTKLEKLMIRLGFTLVLYTVAADVACLFYEQHNRPWEATHNCP CLADLQDQARRPDYAVFMKYFWCLVIGTISGVWWSKLTIESWALCTRCWASKG AAVGAGGSGPGSGPGFPGGGGGGGGGLYSDVSTGLTWRSCTASSVSTPQKPL SQV"

BASE COUNT 403 a 772 c 779 g 467 t

ORIGIN

Query Match 12.1%; Score 89; DB 87; Length 2421;

Best Local Similarity 41.1%; Pred. No. 1.25e-22;

Matches 124; Conservative 47; Mismatches 131; Indels 0; Gaps 0;

Db 306 tgcggtgtgcaaggatcgattacaactacattacattaccccaaccagttcaaccagc 365

| | : | | : | | | | | : | | : | | | | : | | | | : | | |

Qy 140 TNCARYTNTGYCAYGGNATGARTAYCARAAYATGNGYNTCCNAAYYNTYNGNCAYG 199

| | : | | : | | | | | : | | : | | | | : | | | | : | | |

Db 366 acacgaagatgagcgcggttagagtgacacagtttgcgcgtggl-ggagatacagt 425

| | | | | | | | : | | : | | : | | : | | | | | | | | | |

Qy 200 ARACNATGAARGATGTYNNGARCARCGNGNGONTGATNCNTYNTGATGAARCAAT 259

| | | | | | | | : | | : | | : | | : | | | | | | | | | |

Db 426 gctcccggaacctcaagttcttcttgtagcatgtacacgcccacatcgctggagagc 485

| : | | | | | | : | | : | | : | | : | | | | | | | | | |

Qy 260 GYACCCGAYACNNAARAATTTTYNTGWSNYTNTTYGNCNCNGTNTGYTNGAYGY 319

| : | | | | | | : | | : | | : | | : | | | | | | | | | |

Db 486 acaagaagcctctccgccttgctgctctgtgtgacgcgcgaagcgcgctgcgcgc 545

| | : | | : | | | | | : | | : | | | | : | | | | : | | |

Qy 320 TNGAYGARACNATNCARCNTGYCAYWSNYTNTGYTNCARGTNAARGAYMNGTGYCNC 379

| | : | | : | | | | | : | | : | | | | : | | | | : | | |

Db 546 cgcctatgcgcagtaagcgtttgttcgctgaccgatcgctgcgctgcgctgcgcgg 605

| | | | | | | | : | | : | | : | | : | | | | | | | | | |

Qy 380 CNGTNGWSNGNTTYGGNTTYCNGTGGCNGATGYTNGARTGYGAYMNGTTCYCN 439

| | | | | | | | : | | : | | : | | : | | | | | | | | | |

Db 606 ag 607

| | : | | : | | | | | : | | : | | | | : | | | | : | | |

Qy 440 AR 441

RESULT 3 HSU43318 2334 bp mRNA PRI 24-FEB-1996

LOCUS Human putative transmembrane receptor (frizzled 5) mRNA, complete cds.

ACCESSION U43318

NID g1151251

KEYWORDS .

SOURCE human.

ORGANISM Homo sapiens

Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2334)

REFERENCE 1 (bases 1 to 2334)

AUTHORS Wang,Y., Macke,J.P., Abella,B.S., Andreasson,K., Worley,P., Gilbert,D.J., Copeland,N.G., Jenkins,N.A. and Nathans,J.

TITLE A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled

JOURNAL J. Biol. Chem. 271 (8), 4468-4476 (1996)

MEDLINE 96224032

REFERENCE 2 (bases 1 to 2334)

AUTHORS Abella,B., Wang,Y., Macke,J.P. and Nathans,J.

TITLE Direct Submission

JOURNAL Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205

FEATURES

Location/Qualifiers

1..2334

/organism="Homo sapiens"

/chromosome="2"

/map="2q33-34"

/tissue_type="retina"

321..2078

/genes="frizzled 5"

/note="putative transmembrane receptor"

/codon_start=1

/product="transmembrane receptor"

/db_xref="PID:g1151252"

/translation="MARPDPSAPPSLLLLLAQLVGRAAAASKAPVCQEIYPMCRGI GYNLTHMPNQNHDTQDEAGLEVHQFWPLVEIQSPDLRFFLCTMYTPICLPDYHRPL PPCCSVCERAKACGSLMRQYGFAPWRMCRDLPLVGRDAEVLGMDYNRSSEATAPP RPPAKPTLPGPAGSGGCPAGGPFVCKREFVPLIKESHPLYNKRVTGQVPMC AVPCYQPSADERTFATFWIGLSVLCFISTTVALEFLDMDTFRYPERTIIFLSA CYLCSIGFIVRLVGVHASVACSRHNHHYETTGPALELCTIVLLVYFFGNASSLHWV ILSLTFWLAAMKMGNEALAGYQYFHLAHLVPSVKSLTALALSSVDGPVAGICV GNNQNSLRREVLGPLYLLVLTGLFFLAGVFSLFIRSVIKQGGTKTKLEKLMIRI GFTLLYTPASIVACLYEYHVSWEAALTCACPGHDTGQPRAKPEYVWYLMKYF MCLVIGTISGVWWSKLTIESWALCTRCWASKG AAVGAGGSGPGSGPGFPGGGGGGGGGLYSDVSTGLTWRSCTASSVSTPQKPL TGPQPAATYGHKWVSLSHV"

BASE COUNT 355 a 803 c 736 g 439 t

ORIGIN

Query Match 10.2%; Score 75; DB 77; Length 2334;

Best Local Similarity 39.7%; Pred. No. 3.36e-16;

Matches 116; Conservative 44; Mismatches 132; Indels 0; Gaps 0;

Db 439 tctccgc 498

| | | | | | | | : | | : | | : | | : | | | | | | | | | |

Qy 146 TNGCYAGGNATGARTAYCARAAYATGNGYNTCCNAAYYNTYNGNCAYGARACNA 205

| | | | | | | | : | | : | | : | | : | | | | | | | | | |

Db 499 aggcagcagc 558

| | | | | | | | : | | : | | : | | : | | | | | | | | | |

Qy 206 TGAARGARTNTNGARCARGCNGCNGCNGTGGATNCNTYNTGATGAARCAFCAYCAYC 265

| | | | | | | | : | | : | | : | | : | | | | | | | | | |

CDP I R I AM C Q N I G Y N V T K M P L N U G H E L Q T D A E L Q L T T F T P L I Q Y G S S Q Q F L C S Y V
V P M C T E K I N I R I G C G M C L S V K R R E C P F A M P D T I N C V L S K C P P Q N D I N H A K E F
G P C D E E V L P L H K T P I O P G E E C H S V G S N S D E Y I W K R S I N C V L K G Y G D A G L Y S R S A K E F
T D I M M A W A S L C F I S T T F T V T F L I D S S E R S Y P E R T I F L S M C N Y S I A Y I V R L T V G
R E R I S C D F E A A E P L I O B E G L N Q T C A A I F I L M Y F T F F C G M A S T W W V I L T I T F L A A L I G
W G E A T E M S S V E Y I A A A I P A V K T V I L I M R L V D A D D E L T G C V G N Q N L D A L G F W V
A P L F Y I V T G T L F A G L V A L E F I R S N L O G O C T K T D K L E R I M V K I G F S V L Y T P A T C
V I A C Y F Y S N A L F I R S Y A D S N A V M E L K I F M S L V G I T S O M I S A K T I L T F T W Q K S
N I R N S G K V R E K R G N W K P C K G N C T V T *

BASE COUNT	689 a	703 c	711 g	725 t	3 others
ORIGIN					

Query Match 8.6%; Score 63; DB 87; Length 2831;
Best Local Similarity 38.8%; Pred. No. 6.61e-11;
Matches 106; Conservative 40; Mismatches 127; Indels 0; Gaps 0;

Db 515 tgcccaacttagtgggacacgagctgcagacagacgccgagctgcagctgacaaacttca 574

Qy 179 TNCNAYTYTNYTCGNCAYGAPACNATGAARGAGTNYTNGARCARGCGNGCGNTGGA 238

Db 575 cgcgcgtcatccagtagcggtgcgtccagccagctgcagttctcttcttgcggtttatg 634

QY 239 INCGTNGINAIGAAACARTGTCATCCNGAYACNAARAAATTTTINIGWSNIIITIG 299

|| ||| : | | | : ||| : ||| :

[illegible]

359 ARCTAARCAAYMNTGGCCNCCGNTATGNSGCTTTCGTTTCGTCGCCGCAYATGY 418

Db 755 tgaactcagcaagtccccccagaaacgacc 787

419 TNGARTGCGAYMGNTTCCNCARGAYAAYGAY 451

RESULT	6					
LOCUS	DMU6589	2085 bp	DNA	INV	01-SEP-1996	
DEFINITION	Drosophila melanogaster Dfz2 (Dfz2) gene, complete cds.					

ACCESSION U65589
NID 21510050

KEYWORDS fruit fly, source

ORGANISM *Drosophila melanogaster*; Eukaryotae; mitochondrial eukaryotes; Metazoa; Arthropoda;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

AUTHORS Bhanot, P., Brink, M., Samos, C.H., Hsieh, J.C., Wang, Y., Macke, J.E.

TITLE A new member of the frizzled family from *Drosophila* functions as a Windless receptor

JOURNAL Nature 382 (6388), 223-230 (1996)
MEDLINE 96353971

AUTHORS
Bhanot, P., Wang, Y. and Nathans, J.

JOURNAL Submitted (30-JUL-1996) Molecular Biology and Genetics, Johns
 Hopkins University School of Medicine 725 N Wolfe St 805 D
 Baltimore, MD 21205-7089

FEATURES Location/Qualifiers
Baltimore, MD 21205, USA

```
source
1..2083
/organism="Drosophila melanogaster"
```

```

/chrosome="3"
/map="76a"
1..2085
/genes="fz2"
/standard_name="frizzled2"
/notes="Description: frizzled 2; wingless receptor"
/codon_start=1
/product="Dfz2"
/db_xref="PID:g1518051"
/translation="MRNRRLKCVLTILGILLVLTSCRADGPLHSADHGMGCGMGCGRRGLD
ASPAGCYGPAIPKPDNLRCEETIIPMCRGICYNWTSFENWHEUTODEAGLEVFEMPE
PLIVEIKCSDPKKFFLCSMYPTICLEDYHKP L P C V R S V C E A R S G C A P I M Q Y S F E M P E
R M A C E H L P L H G D P D N I L C M E Q P S Y T E A G S G G S G G S G S G K R K Q G G S G S G
C A G S G S S G T C K C R C N S K N C M P O G E K A S K E C S C S C R S P L I F L K E Q I L Q O O S G
P M M H G P H M Y W N L T V Q R I A G V P N C L P C K G O F F S N E K O F A G L I A L M S G L C F C S T I L
T L T T F I D T E F F Y P E R P I V T L S A C Y F M A V G Y L S R N F L Q N E E I A C D G L L R E S S T P
H S C T L F E L L T V F C M A S S I W V I L T T F L A A G L K M G N E A I T K H S O Y F H L A A M L I P T V
Q S V A V L L S A N D G D P I L G I C Y G N I A P D H L K T F V I A P L F V I L I G T T F L M A G F V S L E R
I R S V I Q Q K G G A C A W K A D L K E L I M R I G I S V P T P A T I V G I C Y F E A A Y F E D L I K A
L A C P A Q K G G C K R P L A S Y L K M F Y L A A V G I T S G V I W S K T L E S W R F M R R I L L G A P
D R T C A N Q A L I Q R P P I P H P Y A G S C M G M P V G S A C S L L A T P Y T Q A G G A S V A S T S H H I L
H V I L K Q P A A S H V "
404 a 663 c 612 q 406 t
BASE COUNT

```

	404 a	663 c	612 g	406 t
BASE COUNT				
ORIGIN				

Query Match 8.3%; Score 61; DB 40; Length 2085;
Best Local Similarity 38.5%; Pred. No. 4.75e-10;
Matches 120; Conservative 42; Mismatches 150; Indels 0; Gaps 0;

Db 206 taccaatgttcggggcattggctacaacatgacatccttccccaacgaatgaaccatg 265

Qy 140 TNCARYTNTGCAVGGNATNGARTAYCARAAATATGMGNYTNCNAAYYTNYTNGGNCAYG 199

Db 266 agaccaggacgaagcgggcccggaggtgcaccagtctcggccccctggtggagatcaaat 325

200 ARAGNATGARGARGININGARGCGCGGCGNIGGAINCCNINGINATGNARGCAI 255

[illegible][illegible]

320 TNGAYGARACUATNCARCCNTGTCAYWSNYNTGTCGTCARCTGAARGAYMGTCGGCNC 379

Db 446 ccacatcgcagcagctacagcttcgaatgccggagagaatggcgtcgagcacttgcacc 505

Qy 380 CNGTATGNSGNCNTTYGGNTTYCCNTGGCCNGAYATGYINGARTGYGAYMGNTTYCNC 439

Db 506 ttcatggtgacc 517

Qy 440 ARGAYAYGAYY 451

RESULT 7

10003	AKR03520	4200 bp	mus	mus musculus putative transmembrane receptor (frizzled 7) mRNA
DEFINITION				

U43320
ACCESSION
-1151957
U43320

KEYWORDS
house mouse.

ORGANISM *Mus musculus*
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;


```

RX MEDLINE; 94150718.
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J.,
RA Coulson A., Craxton M., Dear S., Du Z., Durbin R., Faville A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kerehaw J., Kisten J., Laister N.,
RA Latrelle P., Lightning J., Lloyd C., McMurray A., Mortimore B.,
RA O'Callaghan M., Parsons J., Percy C., Rifkin L., Roopra A.,
RA Saunders D., Showkeen R., Smaldon N., Smith A., Sonhammer E.,
RA Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M.,
RA Vaughan K., Waterston R., Watson A., Weinstein L.,
RA Wilkinson-Sproat J., Wohlman P.;
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of
RT C. elegans';
RL Nature 368:32-38 (1994).
CC Current sequence finishing criteria for the C. elegans genome
CC sequencing consortium are that all bases are either sequenced
CC unambiguously on both strands, or on a single strand with both
CC a dye primer and dye terminator reaction, from distinct
CC subclones. Exceptions are indicated by an explicit note.
CC
CC IMPORTANT: This sequence is NOT necessarily the entire insert
CC of clone T23D8. It may be shorter because we only sequence
CC overlapping sections once, or longer because we arrange for a
CC small overlap between neighbouring submissions.
CC The true left end of clone T23D8 is at 1 in this sequence.
CC The true right end of clone T23D8 is at 34576 in this sequence.
CC The true right end of clone T24P1 is at 12749 in this sequence.
CC Coding sequences below are predicted from computer analysis,
CC using the program Genefinder (P. Green, ms in preparation),
CC and other available information.
CC The start of this sequence (1..104) overlaps with the end of
CC sequence CET24D1.
FH Key Location/Qualifiers
FH source 1..34576
FH /organism="Caenorhabditis elegans"
FH /clone="T23D8"
FH /chromosome="I"
FH complement(join(31876..31973,32210..32284,32331..32573,
FH 32640..33032,33192..33264))
FH /product="T23D8.8"
FH /note="protein predicted using Genefinder"
FH /note="Similarity to Mouse bright gene (TR:Q62431)"
FH /db_xref="PID:e304447"
FH /translations="MYELSDVVRKEWLDLWFMHRIKPVTRIPIMAKQVLDLYEL
Y
FT RLVVQHGLVEIINKKLRKITKGLNPSSITSNAFTLTQYLYDYCEKEKLSN
Q
FT SDIQQADIGNRRAPGRRTAPSPFIPQLPHASAAATMLANQLGCMRNDLDDEN
T
FT LSLQASLGFTSTYGAQEWALIEAHQRNLERARAVQVQVARQSIGLTACSNNGNHIH
N
FT SGRSTSSNDSIPAKRPKLENDVKINGASSMRISTKHSNKSYSVSMVEINGITYQ
G
FT VLFAIDTVSES"
FT complement(25597..25908)
FT /product="T23D8.5"
FT /note="protein predicted using Genefinder"
FT /note="Similarity to histone H4"
FT /db_xref="PID:e275927"
FT /translations="MSGRGKGKGLCKGKAKRHKVLRDNTQGTTPAIRLARAGGV
K
FT RISGLIYEETRGVLKVLNVIRDATYCEHAKRKTVTAMDVYALKRQGRITLYGCG

```

```

FT CDS complement(join(7258..7391,7494..7638,8798..8860,9211..934
3
FT ,10019..10101,10150..10212))
FT /product="T23D8.2"
FT /note="protein predicted using Genefinder"
FT /note="Similarity to Rat CD63 antigen (SW:P28648)"
FT /db_xref="PID:e276248"
FT /translations="WVEGGVTIVKYLFLANLVWVGGLSLIVGSILQAKFDNVLDI
L
FT GDERIATPILLIIVIGSLCTLLGFLGCCGAIENYCLTVSFAVLALLITGMWRVHESR
G
FT VESAMDKTHQLFECCGVNSSDWLFTTTPDSCCIEIEGCARENAPLFEFGCHHSVE
Q
FT WVLKNGAMVGICAVLAAIQLVGVCFACCLSKSLKDFHDFY"
FT join(2786..3258,3758..4122,4187..4325,4866..5107,5220..534
FT 0
FT ,5688..5906,6332..6479)
FT /product="T23D8.1"
FT /note="protein predicted using Genefinder"
FT /note="Similarity to Drosophila tissue polarity protein
FT (TR:Q24760)"
FT /db_xref="PID:e77930"
FT /translations="MHRHILILFLGCLSDAQRLSSTSISSMNGFSTTRKCEHITPM
C
FT KNLDYNQVFPNLLGHTTQSEAGPAIAQNPLIKVKCEDIRLFLCTVYVAVCVILEX
P
FT IQPCRELCLSAKNGCESLMKKFGFQWPQLDONKFPVTDLCVGKNSSESNSKKSSND
V
FT TFCVSTIANEVLSPKCPHHMTTSGSHFSLPLLSGRLPECSLTCEADNQPMWFDG
R
FT VRRILRWTAWSVACFVCSILFTLTFTLVDLSRFAPVVRPILYLAFCYLATISTYHIG
V
FT VGEDGFACGTGYSTPTTLVTQGENVGCSALAVVHYFFFMSSCAWLVLCIAWFLAAN
L
FT KWGAESTIALSPYFHAMCGVPALSVTVLVTVNSVDGDTGTCGVGNLNFSAUVFF
F
FT TPIWSTALGAVLLVCGIWSNIRSYIKIQHADVERNISKELEMLRIGAFAMYSL
P
FT TANNAATMWQAVNMPAWLEGMHLHRCVRLQDRELFGFTYPPVDDCPMDPKVAPEIIV
F
FT LIKYVSQLVGCITCAIWWSSKTLSSYKAYIALSRSPTVPAHVQVNMNR"
FT 26277..26660
FT /product="T23D8.6"
FT /note="protein predicted using Genefinder"
FT /note="Similarity to histone H2A"
FT /note="cDNA EST yk72a2.3 comes from this gene"
FT /note="cDNA EST yk124d7.3 comes from this gene"
FT /note="cDNA EST yk124d7.5 comes from this gene"
FT /note="cDNA EST yk72a2.5 comes from this gene"
FT /db_xref="PID:e775250"
FT /translations="MSGRGKGKATGCGAKRSRAGLQPPVGRHLRLILKKNYQAR
V
FT GAGAPVYLAAYLEYLAAEVLELAGNAARDNKKTRIAPRHLQLAVRNDEELNKLACVT
I
FT AOGGVLPNIQAVLLPKKTGGDKR"
FT join(26949..27068,27120..27167,27232..27606,27661..28621,
FT 28689..28819,29049..29748,30273..30464,30510..30659)
FT /product="T23D8.7"
FT /note="protein predicted using Genefinder"
FT /note="cDNA EST yk116f3.3 comes from this gene"
FT /note="cDNA EST yk20f1.5 comes from this gene"
FT /note="cDNA EST yk116f3.5 comes from this gene"
FT /db_xref="PID:e275933"

```

```
FT      /translation="MEDQMLLSAIYDDDLVEKLKVRSSTSRSTSNVPSLENEFLSS
S
FT      SGRSRVDDLPHPIENREPFKLIGKPLSPSTTGRFLSLLANHFQITCNGSIHQYIYR
F
FT      DDPIDPSKKIARTIILRTLEQNPGLIECP LVDGCIHTVYSTELINNVKNSVINVAGV
V
FT      NTKESPNIKLYITHVDSFLDITKIIITGNQDQOKLURMMHAIDTVFRQTSFGHFAVL
Q
FT      SFPSIAQNSAIEPSHGLGMCVTNLGVGREVCYFYQNVVETFDLTMLDVAITTFYR
P
FT      VALVEFLAEILEVPLATVTDGRSLSDVQKKFNFREVAGIKVETRHCSPRFRVARCT
W
FT      KPTENISPHLSETAGQUSKRP LSIUEYKRYNIDITYKHILPCIEVGRTRCILPLEL
C
FT      YVVSQORCIKKLEQQIANLIRATSNATERQNAVMSLQNLRIKMDNDVNAVKGLKVE
A
FT      QLLKTEGRVLPVPRLLYRSPNLKRCQCVTVNNGTWDMRGKGFYSIGIQIREMAIVCFA
S
FT      PELIGEASMSRFRVNLVNVASEIGMPFLEEHRFCRYAEPDQTVKLLHLEHLEQYNLQIV
L
FT      CIVPCKSVVYGELKRRKGLGLATTQCVRSQNVSKASPH T LSNLCKINSKLGGINVIL
S
FT      SPQPSINSEPVLEFIGCHTRSSLASSSDSTSSIAHCDSIACLVGSMGCHPTQSPIF
R
FT      TQPRHQRTIVDMCEMTRAIINFRKSTGTFKHKIIYRAGIADVTVDLIMQTELRAVR
D
FT      ACAMIEYFGPGITFGIDVTHHTRIFAANEKORVGNQSNVPAGTIVETGITVNNLFE
F
FT      YLVSHAGIQGTSRPTKYVVMDDNSIPSIADIEHMTYQLCHTQSRCTRSVSPSVVYA
K
FT      LVAORAKTLMADENFMERFRLCVSHF"
CDS
FT      complement(join(10723..10767,11212..11658,11707..11823,
FT      12477..12917,13249..13431))
FT      /product="T23D8_3"
FT      /note="protein predicted using GeneFinder"
FT      /note="cDNA EST cm16f9 comes from this gene"
FT      /note="cDNA EST CEESF93R comes from this gene"
FT      /note="cDNA EST CEESF93FC comes from this gene"
FT      /db_xref="PID:e304448"
FT      /translation="MGRKAFIDKKSQSFRLVPDNRERSERFKPTQEHLEEQKYGV
Y
FT      YDDYDYDYLQHRVAVNEPMKLVNVEEVEKTTIKTSGSAFPAPPLFLGLVGLKPKPEF
F
FT      DEDVANALEEVTDRNTGELEDNFITLAGLLIDERTTVYSTRRGDSEEDDEDD
E
FT      MYDDYNDDELFGEEAVGEIRVERADQRVIDNAFEELMDREYNTDQIGELDGDYDVG
E
FT      LEFNAGRLHLKINDKGPSNAEYDEELAKHYRMRMLIEEGVINDKEEYEVIEVDEGT
A
FT      ...
Note: remainder of annotations omitted.

Query Match      7.1%; Score 52; DB 8; Length 34576;
Best Local Similarity 40.1%; Pred. No. 2.64e-06;
Matches 120; Conservative 44; Mismatches 130; Indels 5; Gaps 5;

Db      2910 ttccaatgtgcaaatctggattacaacaaacagtatcttccaacatcttcggacata 2969
Qy      140 TNCARYTNTGTCAYGGNATGARTAYCARAAAYATGMYNTGCCNAAYTYNTNGNCAYG 199
Db      2970 caacacaatctgaagtgtgcagcaattgcgaattcaatccattcaatgaagta-aa 3028
Qy      200 ABACNATGAARG- TNYTNGCARGCGNGCGNTGGATGCCNCTNGTNGTNGTNGTNGTNG 258
```

```
Db      3029 tgcgcagaagattcgtctctcttcttctgtactgtctgacacctgtctgtacag-t-ac 3086
Qy      259 TGYCAYCNGCAYACNAAARATTTYTNTGWSNYTNTTYGCNCGCNTGTGYYTNGAYGAY 318
Db      3087 t-cgaaaaaccaatccaaccatgcgagaattgtgtttatctgtcgaataatggatgcgag 3145
Qy      319 YTYAGARACNATNARCCTGTCAYFNSYNTNTGTYTNCARGTNAARGAYMGNTGYGCN 378
Db      3146 tcattaatgaaaaagttgatttcaatggcagatcaatggattgttaacaaattccc 3204
Qy      379 CCNGTATGWSNGNTTYGGNTTYCCNTGGCNGAYATGYTNGARTGYAYMGNTTYCC 437

RESULT 13
LOCUS      CET23D8      34576 bp      DNA      INV      26-NOV-1996
DEFINITION      Caenorhabditis elegans coemid T23D8.
ACCESSION      Z81128
NID      gl628237
KEYWORDS      Caenorhabditis elegans.
ORGANISM      Caenorhabditis elegans
SOURCE      Eukaryote; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabditia; Rhabditidae; Rhabditina; Rhabditidae;
Caenorhabditia.
REFERENCE      1 (bases 1 to 34576)
AUTHORS      Wild,A.
TITLE      Direct Submission
JOURNAL      Submitted (21-OCT-1996) Nematode Sequencing Project, Sanger Centre,
Hinxton, Cambridge CB10 1RQ, England and Department of Genetics,
Washington University, St. Louis, MO 63110, USA. E-mail:
jes@sanger.ac.uk or twenematode.wustl.edu
REFERENCE      2 (bases 1 to 34576)
AUTHORS      Wilson,R., Alnsough,R., Anderson,K., Baynes,C., Berks,M.,
Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J.,
Culson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A.,
Johnson,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M.,
Fulton,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N.,
Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B.,
O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roppra,A.,
Saunders,D., Showkeen,R., Smaildon,N., Smith,A., Sonhammer,E.,
Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M.,
Vaughan,K., Waterston,R., Watson,A., Weinstock,L.,
Wilkinson-Sproat,J. and Wohlman,P.
TITLE      2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans
JOURNAL      Nature 368 (6466), 32-38 (1994)
MEDLINE      94150718
COMMENT      Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
IMPORTANT: This sequence is NOT necessarily the entire insert of
clone T23D8. It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.
The true left end of clone T23D8 is at 1 in this sequence. The true
right end of clone T23D8 is at 34576 in this sequence. The true
right end of clone T24D1 is at 12749 in this sequence. Coding
sequences below are predicted from computer analysis, using the
program GeneFinder (P. Green, ms in preparation), and other
available information.
The start of this sequence (1..104) overlaps with the end of
sequence CET24D1.
```


FEATURES
source

Location/Qualifiers

```
1..34576
/organism="Caenorhabditis elegans"
/clone="T2308"
/chromosome="I"
```

CDS
join(2786..3258,3758..4122,4187..4325,4866..5107,
5220..5340,5688..5906,6332..6479)

/note="protein predicted using GeneFinder"

/codon_start=1

/product="T2308.g"

/db_xref="PID:e275930"

/db_xref="PID:g1628244"

/translation="MHRHLLFLFCLLSADQLSSTSISMNGFSTRKCEHTITPM
CKRLDNGTVFPLNGHTTQSEAPAIQAQNPRLIRKGSIEDIRLFTYAPVGVVLE
KPTQPCRECLSLAKNGCSLMKKEFGQWPDQDCKNFPVTLQVGNSSSSSKSS
NDVTFGVSTIANEVLSPKCPHMHHTTSGSHFLPLLSGLRFLCSLTCEADNQVPM
FQGRVRRILRIWTAMSVACVGSFLTIVTLVDLSRAYPVAPILYLAFCYLAISTV
YMI GVVGDEGFACTYGTPTTLVTQGGENVGCSALAVHVFFFMSSCAWMLVLCIAW
FLANLKGAEISTALSYPHAMCWGPVAVLSVTVLVNVDGVTGTGICSVGNLPS
ALVYFFFTPVVSLAGVALLVCCIMSRIRSYIKLQHADVERNISKEKMLIRICA
FATWSLPTAWAAIMTQAVNMPALWGLHHRVRLQDRLELFGFTYVDDCPMDPK
VAPELIVFLKTVSQLVVGITCAIHWVSSKTLSSVAKYALSSRSPTVPAHVQVN
MR"

CDS
complement(join(7258..7391,7494..7638,8798..8860,

9211..9343,10019..10101,10150..10212))

/note="protein predicted using GeneFinder"

/codon_start=1

/product="T2308.f"

/db_xref="PID:e276248"

/db_xref="PID:g1628243"

/translation="WVEGVTVIKYLLFLANLVLWVGSLIIVGSIILQLKFNVDLI
LGERLATPILLIVIGSLCTLLGFLGCCGAIRENYCLTVSAVLLALLITGMRYHES
RGVSEAWDKYTHQFECQGVNSSDWLTFTTIPDSCSIEIEGCARENAPLFEPGCIHS
VEQVILKNGAWGICAVLAATQVGVCFACCLSKILKQHPFY"

complement(join(10723..10767,10978..11142))

/note="protein predicted using GeneFinder"

/codon_start=1

/product="T2308.e"

/db_xref="PID:e275929"

/db_xref="PID:g1628242"

/translation="MGVTKPFPFKNNKKNISFPLPFFVLVKNKKNKNTTAHQAC
RFFYFKTKMLAKRIGQIKARPI"

complement(join(11188..11658,11707..11823,12477..12917,

13249..13431))

/note="protein predicted using GeneFinder"

/codon_start=1

/product="T2308.d"

/db_xref="PID:e275928"

/db_xref="PID:g1628241"

/translation="MCKRKAFIDKKASQFRLVDPDNTRSERFKPTQEHLEEQKYGV
YDDYDYLAHRAVNEPKLENVHEVEKTTIKTSGSAPPAPPLFLGVALKRPE
FFEDVANALEYTDNRNTEGEDNFIITLAGGLDERTTVYRSTREGDESEEDDDE
DQEDVDDYNDDELFGEEAVGIRVERADQVIONAFEELMDREYNTDQIGELDGDYD
VGGALEPNAGRLHLKLNKQGPSNAEYDEE.AKHYVREMRLLIEGVKQKEEYIEV
DEGTNKKMKIDCESATQYTNIIYHPILTIKEPRGLSRKALKRDKAVEEMDIAEDED
DEDEMEDDDKDESVESTVSTFRPNKNETPQRS LRKKA VKARKLRVVEKANKTPEAE
EKRVKFIHS"

13855..14139

/note="protein predicted using GeneFinder"

/codon_start=1

/product="T2308.h"

/db_xref="PID:e275931"

/db_xref="PID:g1628245"

/translation="MGTIEKNIFRRLVAFEDDDVDADHLETPNFLAFLRVLVLHVLL

CDS

```
GIHQILSAYLLAARSFLETTILDNDHNVQCAAPEVHCFRQEEVCQTS"  
complement(join(13883..14240,14292..15409,15456..16341,  
16423..16656,16705..16807,17668..17721,17774..18278,  
18330..18656,18888..19113,19170..19705,19751..19828,  
19884..19991,20041..20353,21385..21569,21857..21931))
```

/note="protein predicted using GeneFinder"

/codon_start=1

/product="T2308.c"

/db_xref="PID:e276247"

/db_xref="PID:g1628240"

/translation="MHSTGEQRCQAGPYHHMPYHVQDIPPTTSHPSHQRAHPHGLPT
PPVHQQHPNYISMOPPSQOQMTPOALSTQOQOQOQOQOQOQLYSPSPSRGPAAP
AQQKFRTEQVNNQAWPYSQSPAPVPVPSVMSYHQGDDRMSLSVNTTMTNQFPDSQ
RCYSNGSTCENVPENMMQNTIQAEOAIRLMFETPPMORLHMATIRTMIRQDKF
AQVDQANPNQVQIILNIIDYGLKQPVQLPISYYAQLWYNLLDIRRFTLPTISPY
IHQVQMCPRENGQDFRELIQNLISLNWQKOPHKHCANQVQIFNCILMGVKNEX
LRTEQAHLEKLVGTISEYFNPQVHPGMINPAIFIIFREIISKOTRLKDYFIMNN
HSDAIDCFVRSQVCAITTVVQVYNNDLIRACCKLLOVSDAKALAKTPLENLPFL
LRLEIHPDDEVIYSGTGISNVVAHQHVROIAIRSHAFLLHTIISKYPRDELTD
PHQPPPTGLIILKNAMVIGSYRLTAGONPETLPONPELAHLIOVIRTEDLGLLL
APKRRNVCIEICNCLANTANFLAMWIPTPNGETKTAGNEKQVCKFIEIDILKILMS
CLSCGMDPTLLELRSTILRSEFILLRTPFPVKQGVNLVIDENKENLIGHICAAYS
WFRPNNTRTSQQLVERTISLLVMEQCGAEKVAQSYSIDCPILLLNGQV
KPTFIRNVVVCDKTILEHCPTRADIWTDRPMEGLTNRNSDIDAKAANSLSRFEN
DILAAMSREFHAKEDSDSTSSSEDEQKVKSAKFRDLDPMAGPEEDEKRVVRA
QDKKFLDELKGLIKORNAKSNKDLNRLTGDFS LAKAYDKSKTVQRQNVANPRFYI
RSIVEIETYNNKLDKDAKSALSKNNKALPPLRQKLKRYIKDQQLQJLVTDYRVNP
DEGVETPEDEDDGEVSEKAESKPGPSEKAAVSDSDSDSDSDSDSDSNWSSESES
NSSDEDSVTKMEQLRRYFLKKEFRVESKDKKEKKKAVIRKVEAVEEDDAQWTPVN
REKSVHPDPKEEVTHVMKILNEVMSARGKTTDRNQHVANLRKLLEVSSEKEUGL
GINVKIFCIISALFELNAKISDMHEYETETMTTLQTVNSLLDILLIOTRVKLSVTAEE
EDENLKDQTEYRIQCSILLAVORLDGELAKILQNRADGCHNDYIEKLAEKDMSLIE
KAEKVEIRNDSGIFDKHEVQYMMRIEHAYTKYQDQNEEDAGKLMYDLCNKIYTLID
DEKRLQRAMLCHVYLAVDHKWHRARDLLIMSHMQAIVDHSVDVTOQLXNRTICQIG
LCAFRHGFIREAHQGLSEIQNTORAKELLAAQAVCTRQHEKTAQEKIDRSQVPIYMH
INVELMCCVILCSMLLETIPHASCEFEMRRRMLSHSFHYQLKQSEKASLTGPPENR
ENVAASKAMLNGMKKQDYIVNDKMQKYNLFIHNAETVKQWVYRQIESLRTYL
LTVSYVATVSLKLADLELSKQDVHSIISKMIQEEELSATLDEPTDCLIMHRVPS
RQMLALNLDKQLTLAENNEQILEPRTGRGGYQGPSPFPGRNERQKQKQSGGYQ
GERRGCGQDQKRGNGSQGQQRHHPQKRAE"

join(14867..15084,15127..15265)

/note="protein predicted using GeneFinder"

/codon_start=1

/product="T2308.i"

/db_xref="PID:e276249"

/db_xref="PID:g1628246"

/translation="MSFANVTSESAETELADSSVVEDLSVDIRMINDSLHVTHKKKI
SGAVPLWNQIYNMSEHSTLSQSFLIINILLILLIFVQGLDAHHVHFNFVFIEN
ARIVTFNVLLGLFNI"

15774..16091

/note="protein predicted using GeneFinder"

/codon_start=1

/product="T2308.j"

/db_xref="PID:e275932"

/db_xref="PID:g1628247"

/translation="MSNFFLGIKVNNRLLTINWSPIRIIVFLDSFFDSNLSLLFLV
IFALHAEEFFFEQIEPTKLLHLSHTVFIIRRIQFWFRPVRGVIVAVTVYVIAHSCFLRM
FSR"

complement(25597..25908)

/note="protein predicted using GeneFinder"

/codon_start=1

/product="T2308.b"

/db_xref="PID:e275927"

/db_xref="PID:g1628239"


```
/note="protein predicted using GeneFinder; Similarity to
Rat CD63 antigen (SM:P28648)."
/codon_start=1
/product="T2308.2"
/db_xref="PID:e276248"
/db_xref="PID:g1628243"
/translations="MVEGGVTVYKYLFLANLVLVWGGLSLIVGSIQLKFNVLNLDI
LGBERLATPILLVYGLSCLTLGLFGCCAGLRNRYCLTVSAVLLALITGMVRYHES
RGVESAWKTHQLECCGVNSSDMLTFTTIPDSGCCIEIEGCARENAPLEPCCHIS
VEGVKNGAVGGICAVLAIQLVGVCFACCLSKSLIKDFHDFYI"
complement(join(10723..10767,11212..11658,11707..11823,
12477..12917,13249..13431))
/note="protein predicted using GeneFinder; cDNA EST cm16f9
comes from this gene; cDNA EST CEESF93R comes from this
gene; cDNA EST CEESF93FC comes from this gene"
/codon_start=1
/product="T2308.3"
/db_xref="PID:e304448"
/db_xref="PID:g1865750"
/translations="MCKRKAFIDKASQSFRLVPDNRERSERFPTQEHLEEQKYGV
YYDDDYVLOHRAVNEPMKLENVHEVEKTTIKTSGSSAPPAPPLFLGLVALKKPE
FDEVDVANAEEVTDNRNTEGEEDNFITLACGLLDERTVYSTRGEDSEEDDEDE
DDMYDDYNDLFGEEAAGEIRVRADQRVDINAPEELMDREYTDQIGELGGDDYD
VCGALEPNAGRLKHLNUGKPSNAEYDEELAKHYVRERMLIEGVIKDKEEYIEVEV
DECTNKKMKWDCESATQYTNINHYPTLIKREPRGLSRKALKRFDKAVEEMDAEDED
DEDEMEDDDKESFVSFVSTYFNPNETPEQRSKKRAVKAREKRLRVEKKAKTMEAE
EKRLKAGRIQGIKARPI"
complement(join(13883..14240,14292..15409,15456..16341,
18423..18656,16705..16807,17668..17721,17774..18278,
18330..18656,18888..19113,19170..19705,19751..19828,
19884..19991,20041..20214))
/note="Similarity to Yeast nuclear transport protein
(SM:P32497); cDNA EST yk44d4.5 comes from this gene; cDNA
EST yk33b1.5 comes from this gene; cDNA EST yk37d3.5 comes
from this gene; cDNA EST yk8e5.5 comes from this gene;
cDNA EST yk8e7.5 comes from this gene; cDNA EST yk37b1.5
comes from this gene; cDNA EST yk4b7.5 comes from this
gene; cDNA EST yk30g3.5 comes from this gene; cDNA EST
yk8e5.3 comes from this gene; cDNA EST yk8e7.3 comes from
this gene; cDNA EST yk44d4.3 comes from this gene; cDNA
EST yk37b1.3 comes from this gene; cDNA EST yk33b1.3 comes
from this gene; cDNA EST yk37d3.3 comes from this gene"
/codon_start=1
/product="T2308.4"
/db_xref="PID:e304461"
/db_xref="PID:g1865751"
/translations="KSYHQDDRMSLVNTTNTQFPDSQRCYSNGSTCEAVNPEM
MQMNTQAAGQALRWFTNPPDQRLHAKTIRTWIRADQFAQDQANPCVQQLIN
IYDGLQKPQVQLPSYQAQLWYNLLDILRRFTPLIISPYIQVQWQFPCRENGQD
FRELICNLISLNQKQDPMHKCANQVFOIENCIMGVNKEKLRTPEAQHLFEKLVGT
LSEYFNQVPHQMPNPAIFITFRILISKDRLKDYEIWNPNHDPQPPPTGLIILKNA
VMIGSYRLIAGQNPETLPQNPELAHLIQV.IRTFDLLGLLHSDSDIDQFVRSYGVA
ITTVVQVPPNDLIRAGCKLILQVSDAKALAKTPLENIPLFILLRIEIHDPDEVISTGT
GFLSNVNAHQHVKDIAIRSNALFLLHTISKYPRIDELTAPKRRNRVCEIICNCLRT
IANFLVMMHTPTNGETKTAGNEKQVCKFEIDIILKLMSCLSCEGMDTPGLEILRS
TLIRSFILLRLTFVPKQGVLVNIDENKENLIGHICAAYSWFRQPNNTTQSTKQ
LIVERTSILLVLMEOCGAEKVAQYSISIDCPNLNLNGNQKPTF.IHNVLVCKLLE
HCPTRADLWTIDRPMLEGLNHRNSDIKAAANLSRFPENDILAAASRRFFHAKDSO
SDTSSSEDEVQDKVNSAKFRDDLDFMAGPEDEKRVVRAQKDKKFDLAKGIIKQNR
DAKSALSNNAKALPPLRQKLIKKYIKOQQIQDLDVTDVRNPDEGDEYTEDDEDDDFG
EVSESKAKSPGRPSKAAKVSDDSDDDSDSSSEPSNESDSDSDSVTKMEQLRR
YFLKKEFVESKDKKKKRVIRVKEAVEEDDDADMTPVNREKSVVHFDKPEVTHD
VMIKKLEWWSARGKRTDTRNQHVANIRKLLLEVSEKEKELGLGIVNKISFCIISALFEL
```

CDS

CDS

```
NAKISHMEYETFTTIQTWNSLLDLLIGTDVRKLSVITYAEEDENLKDDTQETRIQGS
ILIAVORLDCELAKLQNAOCHSDNYIEKLKAEKWCMSLIEKAEKVELRNDSGIFDK
HEVCVYMMRIEHAHYKYQDQNEEDACKLMDYLCKNKIYTLDDDEKRLRQRAMLCHVYVL
AVDKMHRARDLLLASHMQAIYDHSDDVTQILLYNRTIQUGICAFRHG IREAHQGLS
EIQNTQKARDELLAQAVGTQRQHTAEQEKIDRSQRPYHMHINVELMCVCTJCSMLL
EIPHASCEFEARRMLRSRSHYQLQKSEKASITGPPENTREHVAASKAMLNDGDKK
CQDIYVNDKQNVNLFHNAETVAGWVRRIOESLATYLLRYSYTVATVYLSLKLAD
LFELSKDVHSITISKMIIGQELSATIDEPDTCLIMHRVPSRLQMLALNSDKLQTLA
ENNEOILPRTGRGGYQCGSWFPPGRNERQDKQKSGSGYQERRGGQCGQDQKRGWG
SGGQQRHREQPRAF"
complement(25597..25908)
/note="protein predicted using GeneFinder; Similarity to
histone H4"
/codon_start=1
/product="T23D8.5"
/db_xref="PID:e275927"
/db_xref="PID:g1628239"
/translations="MSGRKGKGLGKGGAKRRHKVLRDNIQGITKPAIRRLARRGGV
KRISGLIYEETRGVLK-VLENVIRAVTYCEHAKRKVTYANDVVVYALKRQGRITLYGEG
G"
26277..26660
/note="protein predicted using GeneFinder; Similarity to
histone H2A; cDNA EST yk72a2.3 comes from this gene; cDNA
EST yk124d7.3 comes from this gene; cDNA EST yk124d7.5
comes from this gene; cDNA EST yk72a2.5 comes from this
gene"
/codon_start=1
/product="T23D8.6"
/db_xref="PID:e276250"
/db_xref="PID:g1628248"
/translations="MSGRKGKGAKTGGKAKSSRSRAGIQPVGRLHRLIRLKGNYAQR
VGAQPVYLAUVLEVLAELVLEIAGNARUNKTRIAPRHQLQAVNDEELNKLKLAGV
TAQGGQVNLQAVLLPKTGGDKK"
join(26949..27068,27120..27187,27232..27606,27661..28621,
28689..28819,29049..29748,30273..30464,30510..30659)
/note="protein predicted using GeneFinder; cDNA EST
yk116f3.3 comes from this gene; cDNA EST yk20f1.5 comes
from this gene; cDNA EST yk116f3.5 comes from this gene"
/codon_start=1
/product="T23D8.7"
/db_xref="PID:e275933"
/db_xref="PID:g1628249"
/translations="MEDQMLLSAIYDDDLVEKLKVRSSSTSRSTSINVPSENEFLSS
SSGSRVSDLLYLRPTEENREPFKLIGRP.LPSTTGRFLSLLANHFOITCNGSIIHOYI
RFPDPIPKKLIARTILRLQEQNPGLIECP.LVPDGLHVTYSTELINVEKVNNSVINVA
GVNTPKSPNLFKLYLTHVDSFLDITKITQNDQNKLRMHAIDTVFRGTSTGNFH
AVLOSFS.IAONSAILP-HGLGMGTVNLGVREVCYGYQNVVETFDLTWNLDVAAT
TYFRVPAVEFLAEILEVLPATVTDGSRSLDVQKKNREVAGLKVEYTHSCSPRFR
VARCTNKPTEISF.H.SETAGNQDSKPLSLIVEYKRRYNIIDLTKHLPCEIENVTRREC
ILPLELCVYVSGQRCKIKNEQILANLRATSRNATERQNAVMSLQNLKMDNDVNAV
KGLGVCAQELKITEGRLVPLRLLYNSPLKROCVTPVNNGTWDRGNKFNYSICQIR
ENAIYVCAPEITIGASMSRFVNLVNVASEICMPFLEEHRFCRYAEPDQTKLLIHL
NQYNLQILVLCVPGKSVVYCELRKKGELLGLTTCVRSQNVSKASPHPTLSNLMKIN
SKLGINVILSSPPQSINSEPVLETCGLHTRSSSLASSDSTSSIAHCOSSTACIACGSM
DGHPTQSP.IFRTHQRTIDVDMCEMTREAIINFRKSTGFKPKHIIYRAGIADTVV
DEIMQTELRAVRDA.CAMIEYGFQPGTIFGLDVTHTHTRLEAANEKDRVGNQNPVAGT
LVETGIVNNLFEELVNSHAGIQGTSRPTYVVMDDNSIP.SAD.IHENTYQLCHTQSR
CTRSVSPSPVYAKLVAQRAKILMADENDFMERFLCVSHF"
complement(join(31876..31973,32210..32284,32331..32573),
...
```

Note: remainder of annotations omitted.

Query Match 71.1%; Score 52; DB 111; Length 34576;
Best Local Similarity 40.1%; Pred. No. 2.64e-06;

Dec 9 10:15

848439-3-trans.rng

1

WVQSEH (TW)

Release 2.10 John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MParch_mn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 10:11:03 1997 MasPar time 92.45 Seconds
931.301 Million cell updates/sec

Tabular output not generated.

Title: >848439-3-trans

Description: (1-825) from translate.seq

Perfect Score: 688

N.A. Sequence: 1 WNGCNMGNGNTTYT.....SNATNGNAARYTNCARTGY 825
Comp: SWNGCNKCNCRANARRA.....WNTANKNTTYRANGTYACR

Scoring table:

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 142080 seqs, 52183452 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: n-geneseq28

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 11.254; Variance 12.824; scale 0.878

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	122	17.7	1047	2	Human Natriuretic Pep	3.70e-29
2	99	14.4	1047	2	Human Natriuretic Pep	3.46e-21
3	53	7.7	501	3	Sequence encoding new	5.08e-06
4	53	7.7	501	3	Sequence encoding new	5.08e-06
5	53	7.7	565	6	HCV envelope region n	5.08e-06
6	52	7.6	501	3	Sequence encoding new	1.04e-05

Dec 9 10:15

848439-3-trans.rng

2

c	7	51	7.4	498	3	N50034	Sequence encoding new	2.11e-05
c	8	51	7.4	501	3	N50023	Sequence encoding new	2.11e-05
c	9	51	7.4	501	3	N50028	Sequence encoding new	2.11e-05
c	10	50	7.3	501	3	N50031	Sequence encoding new	4.27e-05
c	11	50	7.3	501	3	N50023	Sequence encoding new	4.27e-05
c	12	50	7.3	501	3	N50029	Sequence encoding new	4.27e-05
c	13	49	7.1	501	3	N50024	Sequence encoding new	8.62e-05
c	14	49	7.1	501	3	N50032	Sequence encoding new	8.62e-05
c	15	49	7.1	501	3	N50025	Sequence encoding new	8.62e-05
c	16	49	7.1	501	3	N50024	Sequence encoding new	8.62e-05
c	17	49	7.1	984	17	Q94336	Degenerate Alteromona	1.74e-04
c	18	48	7.0	501	3	N50026	Sequence encoding new	3.48e-04
c	19	47	6.8	498	3	N50034	Sequence encoding new	3.48e-04
c	20	47	6.8	501	3	N50031	Sequence encoding new	3.48e-04
c	21	47	6.8	501	3	N50029	Sequence encoding new	3.48e-04
c	22	47	6.8	501	3	N50032	Sequence encoding new	3.48e-04
c	23	47	6.8	501	3	N50028	Sequence encoding new	3.48e-04
c	24	45	6.5	501	3	N50027	Sequence encoding new	1.38e-03
c	25	45	6.5	501	3	N50030	Sequence encoding new	1.38e-03
c	26	45	6.5	501	3	N50033	Sequence encoding new	1.38e-03
c	27	45	6.5	565	6	Q35072	HCV envelope region n	1.38e-03
c	28	43	6.2	501	3	N50026	Sequence encoding new	5.42e-03
c	29	42	6.1	204	1	N81164	Base substituted E.co	1.06e-02
c	30	41	6.0	204	1	N81164	Base substituted E.co	2.08e-02
c	31	41	6.0	657	7	Q43519	Degenerate FMN reduct	2.08e-02
c	32	40	5.8	91	9	Q51746	Oligonucleotide probe	4.05e-02
c	33	39	5.7	91	9	Q51746	Oligonucleotide probe	7.85e-02
c	34	39	5.7	501	3	N50025	Sequence encoding new	7.85e-02
c	35	39	5.7	774	3	Q22981	Gelonin toxin DNA	7.85e-02
c	36	39	5.7	984	17	Q94336	Degenerate Alteromona	7.85e-02
c	37	39	5.7	1065	17	Q4335	Degenerate Alteromona	7.85e-02
c	38	39	5.7	3871	2	N71302	HSV-1 gB and surround	7.85e-02
c	39	39	5.7	3871	2	N71302	HSV-1 gB and surround	7.85e-02
c	40	38	5.5	657	7	Q43519	Degenerate FMN reduct	1.51e-01
c	41	37	5.4	711	15	Q88282	V.fischeri flavin red	2.89e-01
c	42	37	5.4	1065	17	Q94335	Degenerate Alteromona	2.89e-01
c	43	35	5.1	420	14	Q88233	Bovine aFGF synthetic	1.04e+00
c	44	34	4.9	711	15	Q88282	V.fischeri flavin red	1.96e+00
c	45	33	4.8	997	3	Q22702	Sequence encoding Pla	3.66e+00

ALIGNMENTS

RESULT 1
ID Q10572 standard; DNA; 1047 BP.
AC Q10572;
DT 09-APR-1991 (first entry)
DE Human Natriuretic Peptide Receptor B.
KW PFRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
OS hyperaldosteronism; glaucoma; guanyl cyclase.
FH Homo sapiens.
FT Key Location/Qualifiers
FT Peptide 1..22
FT /label= signal sequence
FT Protein 12
FT /label= mature NPBR
FT Domain 23..455
FT /label= extracellular domain
FT /note= "binds natriuretic peptides A,B and C"
FT Domain 456..456
FT /label= transmembrane domain
FT Domain 479..1047
FT /label= cytoplasmic domain
FT /note= "GC and protein kinase activity"

FT	Modified -site 24..26
FT	/label= N-glycos _site
FT	Modified -site 35..37
FT	/label= N-glycos _site
FT	Modified -site 161..163
FT	/label= N-glycos _site
FT	Modified -site 195..197
FT	/label= N-glycos _site
FT	Modified -site 244..246
FT	/label= N-glycos _site
FT	Modified -site 277..279
FT	/label= N-glycos _site
FT	Modified -site 349..351
FT	/label= N-glycos _site
FT	Modified -site 600..602
PN	W09100292-A.
PD	10-JAN-1991.
PF	22-JUN-1990; U03586.
PR	23-JUN-1989; US-370673.
PA	(GETH) GENENTECH INC.
PI	Chang M, Goeddel D, Lowe D;
PJ	WPI; 91-036711/05.
DR	N-PSDB; Q10324.
PT	Natriuretic protein receptor B - for diagnosis and treatment of
PT	kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS	Claim 3; Fig 1; 49pp; English.
CC	The sequence was derived from the DNA encoding natriuretic peptide
CC	receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC	kinase activity. The DNA can be inserted into expression vectors
CC	for the prodn. of the protein, opt. after being mutated to produce
CC	NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=
CC	114,952). The protein (or variants) can be used in treatment of
CC	natriuretic peptide disorders, and also to isolate peptides using
CC	affinity chromatography. Antibodies with affinity for NPRB can
CC	also be prepd.
SQ	Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;
Query Match 17.7%; Score 122; DB 2; Length 1047;	
Best Local Similarity 12.3%; Pred. No. 3.70e-297;	
Matches 100; Conservative 230; Mismatches 477; Indels 8; Gaps 8;	
Df	177 drhyntngvnnanngsnshvnyarngngnnathnnrzngrnvynccgnmmnhnn 236 ::: : : : : : : : : : : : : : : : : : :
Qy	10 GGNYTNTTYTTTGNCARCNGCAVTTWSNTAYAAARMGNSAAYTGYAACCNATN 69 : : : : : : : : : : : : : : : : : : : :
Df	237 mnanrnrtndgdyvnyndvngnszagnratrgtwnrdhttrnnanznaantvnvn 296 :
Qy	70 CCNGCWAAYTN-CARYTTGTGTCAGGNATNGARTACARAAATATGMGNYTCCNAAYT 128 :
Df	297 tyrrnnnnnnnnnrnfardndgvyngnsmnmnagcnydgqnnnyvanvntntnnng 356 :
Qy	129 NYTNGNCAYGARACNATGAARGA-RGTNTYTGACRARGCGNGCTGGATWCQNYNG 187 :
Df	357 trndgrnrvkmgrgyhgtvgtnvmckndrntdvnmwmgdndsgdnnaahysgank 416 :
Qy	188 TNATGAACARTGYTCAYCCGACACAAARAARTTYTNTGYWSNYTNTGYGCCNGCNT 247 :
Df	417 nmwtgrnrvnwkgansdnncandndsdcdktstnatnaavngtgnntmngvsenn 476 :
Qy	248 GYTTCAGCATGYTCAGACNATCARCCTGTGYCAYSNYTNTGYTCNCARGTVAARG 307 :
Df	477 nnrknnmnknmasmw-rnfrwmnnngnsnrhkgagerntnrgssygenmtahgkyn 535 : : : : : : : : : : : : : : : : : : : :

Qy	308	AYMWTGCGNCNCNGTATGSHSGNCITTYGGNTTYCNTGCGCCGCAVATGYTNGARTGCG	36
Db	536	nnantghnkvgnvankhvkrrntrnrvnnnkhmrdivnnhntrnngaendnnnnncv	595
Qy	368	AYMGN-TTYCNCARGAAYAYGYTNTGYATNCYNTGNCWNSGNSGAYCAYTYTNTN	426
Db	596	tnycrgsnidnnndnnndwmrynnnndvkvgmnnhnnsehgenskcnsvdtr	655
Qy	427	CCNGCNCARGCAGCCNCAARGTNTGCG-ARGCNGTYAARAAYAAARAAYGAYGAYGA	485
Db	656	nvntdygnasrstannddnnyakknntannnsgnnnttgmnaadvysngnnnn	715
Qy	486	YAAYG-AVATNATGARGACNVTNTGTYAARAAYGATTYTCGYNTNAAARATNAARGTAARG	544
Db	716	nnazsgnynngndgnkhnnkvngnrgnrynnrnsndrtnnnnnnvnmrcwandn	775
Qy	545	ARATNACTATATNAAAYMCGNAGYACNAARATNATYTNCGARACNAARWSNAARACNATNT	604
Db	776	rndngnkvgnrrnnknnggtsemdnnnrmmnyannnnkvnnrtnaynnnkrkanmy	835
Qy	605	AYAAPTYNAAAYGGNTGNSGARGMGAGYTYNAAARARWSNGTYTNTGGYTN-AARGAY	663
Db	836	nnnnhsevanmnkrgntvnanandsvtnysndvngntaneanstmmvtnnnndnytcnda	895
Qy	664	WSNTYTCARTGYACNTGYCARGARATCAAYGAYATNAAAYGCNCNTAYTNTGNTATGCGN	723
Db	896	ndndndykvntngdaynvvevgngnrgnrrhannmarmanandavesnrnrhrndnr	955
Qy	724	CAARACARGCN-CCNGARYTNGTNTATNACNWSNGTNAARMGNTGGCABAARGCGNARMG	782
Db	956	nrngvhtgncvavgvgnkmnrycnngdvtncasm	990
Qy	783	NGARTTYAARMGNATNWSNMGNSNATNMGNAARY	817
RESULT 2			
ID	Q10572	standard; DNA; 1047 BP.	
AC	Q10572;		
DE	09-APR-1991	(first entry)	
DT	Human Natriuretic Peptide Receptor B.		
NC	NPBB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;		
KW	hypertalosteronism; glaucoma; guanyl cyclase.		
OS	Homo sapiens.		
FS	Key	Location/Qualifiers	
FT	Peptide	1..22	
FT	/label=	signal_sequence	
FT	Protein	12	
FT	/label=	mature_NPBR	
FT	Domain	23..455	
FT	/label=	extracellular domain	
FT	/note=	"binds natriuretic peptides A, B and C"	
FT	Domain	456..456	
FT	/label=	transmembrane domain	
FT	Domain	479..1047	
FT	/label=	cytoplasmic domain	
FT	/note=	"3C and protien kinase activity"	
FT	Modified -site	24..26	
FT	/label=	N-glycos site	
FT	Modified -site	35..37	
FT	/label=	N-glycos site	
FT	Modified -site	161..163	
FT	/label=	N-glycos'site	
FT	Modified -site	195..197	
FT	/label=	N-glycos site	

FT	Modified -site	244..246	
FT	/label= N-glycosite		
FT	Modified -site	277..279	
FT	/label= N-glycosite		
FT	Modified -site	349..351	
FT	/label= N-glycosite		
FT	Modified -site	600..602	
FT	/label= N-glycosite		
PN	W09100292-A.		
PD	10-JAN-1991.		
PD	22-JUN-1990; U03586.		
PR	23-JUN-1989; US-370673.		
PA	(GETH) GENENTECH INC.		
PI	Chang M, Goeddel D, Lowe D;		
PI	WPI; 91-036711/05.		
DR	N-PSDB; Q10324.		
PT	Natriuretic protein receptor B - for diagnosis and treatment of		
PT	kidney failure, heart failure, hypotaldosteronism, glaucoma etc.		
PS	Claim 3; Fig 1; 49pp; English.		
CC	The sequence was derived from the DNA encoding natriuretic peptide		
CC	receptor B, NPRB, having quanyl cyclase (GC) activity and protein		
CC	kinase activity. The DNA can be inserted into expression vectors		
CC	for the prodn. of the protein, opt. after being mutated to produce		
CC	NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=		
CC	114,952). The protein (or variante) can be used in treatment of		
CC	natriuretic peptide disorders, and also to isolate peptides using		
CC	affinity chromatography. Antibodies with affinity for NPRB can		
CC	also be prepd.		
SQ	Sequence	1047 BP; 87 A; 15 C; 83 G; 51 T;	
	Query Match	14.4%; Score 99; DB 2; Length 1047;	
	Best Local Similarity	12.4%; Pred. No. 3.46e-21;	
	Matches	81; Conservative 178; Mismatches 392; Indels 2; Gaps	

[illegible]

Cp	262	CNARTCTCTNARBCACNCGCCCAANARWSCRANARRAYTTTNGTCTCGRT	203
Db	852	vnanandsvtnydsnvgntaneastnrv-vtnmndyctendanndvdykvtngd	910
Cp	202	GRCAYGTGTTTCATNACNARNGGNATCCANGCNCNGCYGYTCNARNACTYCTTCATNG	143
Db	911	aymvsgngngnrgnrbhannnarmnandavssnrnrhndnrrnngvhtgnvcagvv	970
Cp	142	TYTCRTGNCNARNARRTTNGSNARNKCATRTTYTGATYTCNATNCCRTGCANARYT	83
Db	971	gnkmzryngdvtntasrmnsngnankhsvsttkdandngcnmmnrgdv	1023
Cp	82	GNARTTNGCNGGNATNGYTCRTCARCTTNWSCKYTTTANWSRAARTCNGGY	30
RESULT	3		
ID	N50030	standard; DNA; 501 BP.	
AC	N50030;		
DE	04-SEP-1991	(first entry)	
DT	Sequence encoding new modified human beta interferon polypeptides		
DE	IFX 447.		
KW	Antiviral; cell growth regulator; immune system regulator;		
KW	antiproliferative; ss.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	1..501	
FT	/*tag= a		
PN	EP-163993-A.		
PD	11-DEC-1985.		
PF	17-MAY-1985; 105750.		
PR	17-MAY-1984; GB-012564.		
PA	(SEAR) SEARLE G D & CO.		
PI	Bell LD, Boseley PC, Porter AG;		
PI	WPI; 85-311944/50.		
DR	P-PSDB; P50029.		
PT	New modified human beta interferon polypeptide(s) - prepd. by		
PT	plasmid transformed bacteria, with improved antiviral,		
PT	anti-proliferative and immune regulating actions		
PS	Claim 28; Chart 2h, page 39; 71pp; English.		
CC	INF of the invention are more active and have different affinities		
CC	for cell surface receptors (allowing selective targeting); they		
CC	have higher therapeutic index; improved stability against microbial		
CC	breakdown during synthesis; and better in vivo solubility and		
CC	stability. They are also easier to recover from incubation mixts.		
SQ	sequence 501.BP; 110 A; 31 C; 69 G; 79 T;		

[illegible]

Cp 343 GBAACCCBRANCKNSCATNACNGCGRCAKCTCYTTNACYTGNACRANARNSRT 284
 Db 250 gngaraaytbytbgnaaaygntaycaycarathaaycaytbaaraacngntytbgar 309
 Cp 283 GRCANGCYTGATNGTTCCTCNARRCTCRNARRCANACNGCGRANARNSRANA 224
 Db 310 garaarytbqaraargaytytvaacnmddgnaarytbatgwnwnnytbcaaytbaar 369
 Cp 223 RRAAYTYTTTNGTTCNGRTGRCAYTYTTCATNACNARNGNATCANGCNCNGCYT 164
 Db 370 mgdtaaytvggmddgthybcaaytayaabaargnaatgartaaywnncaytygcntgg 429
 Cp 163 GYTCAARNACYTCTTCATNGTTCRTGNCNARRTTNGCNGNARNCCKAYRTTYTGT 104
 Db 430 acnathgtnmddgngarathymbgdaaytytytayaathaaymgdybaacngntay 489
 Cp 103 AYTCAATNCRTGRCANARYTGNAARRTTNGCNGNATNGCYTTCARTTWSNCKYTRT 44
 Db 490 ytbmgdaay 498
 Cp 43 ANWSRAART 35

RESULT 4

ID N50027 standard; DNA; 501 BP.
 AC N50027;
 DT 04-SEP-1991 (first entry)
 DE Sequence encoding new modified human beta interferon polypeptides
 DE IFNX 444.
 KW Antiviral; cell growth regulator; immune system regulator;
 KW antiproliferative; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..501
 FT /*tag= a
 PN EP-163993-A.
 PD 11-DEC-1985.
 PF 17-MAY-1985; 105750.
 PR 17-MAY-1984; GB-012564.
 PA (SEAR) SEARLE G D & CO.
 PI Bell LD, Boseley PG, Porter AG;
 DR WPI; 85-311944/50.
 DR P-PSDB; P50026.
 PT New modified human beta interferon polypeptide(s) - prep'd. by
 PT plasmid transformed bacteria, with improved antiviral,
 PT anti-proliferative and immune regulating actions
 PS Claim 28; Chart 2e, page 36; 71pp; English.
 CC Compared with interferon beta prep'd. by recombinant methods, the
 CC INEs of the invention are more active and have different affinities
 CC for cell surface receptors (allowing selective targeting); they
 CC have higher therapeutic index; improved stability against microbial
 CC breakdown during synthesis; and better in vivo solubility and
 CC stability. They are also easier to recover from incubation mixts.
 SQ Sequence 501 BP; 112 A; 31 C; 67 G; 80 T;

Query Match 7.7%; Score 53; DB 3; Length 501;
 Best Local Similarity 16.0%; Pred. No. 5.08e-06;
 Matches 72; Conservative 135; Mismatches 240; Indels 4; Gaps 3;
 Db 53 araarytbyttbgacarybaaygmgdytbgarataygytbaargaymgdatgaayt 112
 Cp 668 ARNWSRTCTTNARCCANARNAWSYTTTNNARRTCNCKYTCNWSNACNCCRTTARY 609

Db 113 tyggnathcngargarathaaarabytbcacarttycataargargaygngcnybta 172
 Cp 608 TTRANATNGYTTNWSYTTNGTTCNARNATNATYTTNGTTCNCRATTNATRTANGTN 549
 Db 173 cnathayaratgytbcaayathtygcnathhtymgdcargaywnwnwnnacng 232
 Cp 548 ATYCYTTNACYTTNATYTTNARNGCRAARTCTRTTTCANARNGTYTCCATNATYTCR 489
 Db 233 gntggaaygaracnathgngaraaytytbytbgnaaaygntaycaycarathaaycayy 292
 Cp 488 TTRTCRTCTRTTTRTTRTTCANGCYTCRANACYTTNGGCGYTCYTCNGTNGCN 429
 Db 293 tbaaraengnytbgarara--arytbgaraa--rgargaytytvaacnmddgnaarytba 349
 Cp 428 GGNARNARRTGTCTNWSNCGNARNGNATRCANARRCTRTTTCYTCNGGGRANCKR 369
 Db 350 tgnwnwnnytbcaaytbaarmgmtaytvggmddgthybcaaytayaabaargnaarg 409
 Cp 368 TCRCAYTCAACATRTGCGCGCANGGRAANCRAANCNWSATNACNGCGRCAACR 309
 Db 410 artaywnncaytygcntggacnathgtnmddgngarathymbgdaaytytytaya 469
 Cp 308 TCYTTNACYTGNACRANARNSRTGRCANGCYTGATNGTTCRTCTNARRTCNARR 249
 Db 470 thaaymgdybaacngntay-ytbmgdaayt 499
 Cp 248 CANACNGCGRANARNAWSRCAARRAAYT 218

RESULT 5

ID Q35072 standard; DNA; 565 BP.
 AC Q35072;
 DT 20-MAY-1993 (first entry)
 DE HCV envelope region nucleic acid.
 KW Envelope; region; type C; hepatitis; virus; HCV; vaccine; serum;
 KW non-A, non-B; amplify; ss.
 OS Hepatitis C virus.
 PN J04349885-A.
 PD 04-DEC-1992.
 PF 29-MAY-1991; 152169.
 PR 29-MAY-1991; JP-152169.
 PA (TEIJ) TEIJIN LTD.
 DR WPI; 93-022708/03.
 PT Envelope region nucleic acid fragment - for type C hepatitis
 PT virus (1), for producing vaccine
 PS Claim 1; Page 2; 13pp; Japanese.
 CC This sequence encodes a novel envelope region of type C hepatitis
 CC virus (HCV). This fragment can be used for the preparation of a
 CC vaccine for hepatitis C. This fragment was prepared from the serum
 CC of non-A, non-B hepatitis patients and the envelope region DNA was
 CC amplified by PCR using the primer sequences given in Q35073-76.
 SQ Sequence 565 BP; 61 A; 92 C; 106 G; 85 T;

Query Match 7.7%; Score 53; DB 6; Length 565;
 Best Local Similarity 13.7%; Pred. No. 5.08e-06;
 Matches 61; Conservative 160; Mismatches 218; Indels 5; Gaps 5;

Db 60 yttdecygtggytytctatctcttytgytgytbytbtytgytgytgytgytgyt 119
 Cp 1 WSNCGNWSGNGYNTTYTNTTYTGGNCARCNGAYTWSNTAYARNGNWSNAAATGY 60
 Db 120 hgytctctgytgytgytgcrcacacsswcmggnnyrtaycatgtcacba-rydaytgy 178
 Cp 61 NARCCNATNCCNCAARNAAYTTCARYTNTCYCAGCNATNGARTAYCARAATGNGNTN 120

Best Local Similarity 8.0%; Pred. No: 2:11e-05;
Matches 39; Conservative 181; Mismatches 265;

[illegible]

RESULT	8	
ID	N50023	standard; DNA; 501 BP.
AC	N50023;	
DT	04-SEP-1991	(first entry)
DE	Sequence encoding new modified human beta interferon polypeptides	
DE	IFNX 416.	
KW	Antiviral; cell growth regulator; immune system regulator;	
DE	antiproliferative; ss.	
OS	Homo sapiens.	
FT	Key	Location/Qualifiers
CDS	FT	1..501
FT	/*tag= a	
PN	EP-163993-A.	
PD	11-DEC-1985.	
PF	17-MAY-1985; 105750.	
PR	17-MAY-1984; GB-012564.	
PI	(SEAR.) SEARIE G D & CO.	
PA	Bell LD, Boseley PG, Porter AG;	
DR	WPI; 85-311944/50.	
DR	P-PSDB; P50022.	
PT	New modified human beta interferon polypeptide(s) - prepd. by	
PT	plasmid transformed bacteria, with improved antiviral,	
PT	anti-proliferative and immune regulating actions	
PS	Claim 28; Chart 2a, page 32; 71pp; English.	

Compared with interferon beta prepd, by recombinant methods, the INFs for cell surface receptors (allowing selective targeting); they have higher therapeutic index; improved stability against microbial breakdown during synthesis; and better in vivo solubility and stability. They are also easier to recover from incubation mixts. 501 BP: 107 A: 31 C: 69 G: 80 T:

	Query Match	7.4%;	Score 51;	DB 3;	Length 501;
	Best Local Similarity	16.0%;	Pred. No. 2,11e-05;		
	Matches	50;	Conservative	95;	Mismatches 167; Indels 1; Gaps 1;
Db	166	gcnymbacnathaygaratgytbcaraayathtygcgnathttymgdmcargaywnnnw	225		
	:	: : : : : :	: : :	:	:
Qy	410	SWSNGAYCAYTYNTYNCNGCNACGARGCCNCAARGTNTGYGARGCNTGYAARA	469		
	:	: : : : : :	: : :	:	:
Db	226	wnacngngntgaaygaracnathgtngaraaytbytbqcnaaaygtntaycaycarat	284		
	:	: : : :	: : : : :	:	:
Qy	470	AYAARAAGYGAAYGAYATNATGARACNYTNTGYAARAAYGAYTTYGCNYTNA	529		
	:	: : : : : :	: : : : :	:	:
Db	285	haaycyytbaaracngntycbgargataarytbgaraargarayttycmddgnaa	344		
	:	: : : : : : :	: : : : :	:	:
Qy	530	ARATNAARTNAARGARATNACNTAYATANAAVMGNMGAYACNAARATNATNTYNGARACNA	589		
	:	: : : : : :	: : : : :	:	:
Db	345	rytbatgwinnwytcayytbaarmgdtaytaygcmgdathytcbaytcaaytbaargc	404		
	:	: : : : : : :	: : : : :	:	:
Qy	590	ARWSNAARCANWNTAYARYTNAAVYCGNGTWNSGARMGCGAYYTNAABAARWSNCTNY	649		
	:	: : : : : :	: : : : :	:	:
Db	405	naargattaywnncaytygcntggagcnathgtmmgdgtngarathybtmgdaayttyta	464		
	:	: : : : : :	: : : : :	:	:
Qy	650	TNTGGYTNAARGAWSNYTCARTGYACNTGYGARGARATGAAYGAYATNAAYGCNCNT	709		
	:	: : : : : :	: : : : :	:	:
Db	465	ytythaaymgd	477		
	:	: : : : :	: : : :	:	:
Qy	710	AYYTGNTATGGG	722		

RESULT	9	
ID	N50028	standard; DNA; 501 BP.
AC	N50028;	
DT	04-SEP-1991	(first entry)
DE	DE	Sequence encoding new modified human beta interferon polypeptides
DE	IFNX 445.	
KW	Antiviral;	cell growth regulator; immune system regulator;
KW	antiproliferative;	ss.
OS	Homo sapiens.	
Key		Location/Qualifiers
FT	CDS	1..501
FT	/*tag= a	
PN	EP-163993-A.	
PD	11-DEC-1985.	
PF	17-MAY-1985;	105750.
PR	17-MAY-1984;	GB-012564.
PI	(SEAR)	SEARLE G D & CO.
PI	Bell LD, Boseley PG,	Porter AG;
DR	WPI;	85-311944/50.
DR	P-PSDB;	P50027.
PT	New modified human beta interferon polypeptide(s) -	prepd. by
PT	plasmid transformed bacteria, with improved antiviral,	
PT	anti-proliferative and immune regulating actions	
PS	Claim 28; Chart 2f,	page 37; 71pp; English.
CC	Compared with interferon beta prep.	by recombinant methods, the
CC	INRs of the invention are more active and have different affinities	
CC	for cell surface receptors (allowing selective targeting);	they

CC have higher therapeutic index; improved stability against microbial CC breakdown during synthesis; and better in vivo solubility and CC stability. They are also easier to recover from incubation mixts. Sequence 501 BP; 114 A; 31 C; 68 G; 79 T;

Query Match 7.4%; Score 51; DB 3; Length 501;
Best Local Similarity 9.6%; Pred. No. 2.11e-05;
Matches 44; Conservative 159; Mismatches 252; Indels 2; Gaps 1;

Db 2 tgmntayaaytbybtbggnttytbcargdwnnnaaytctgcaraatytby 61
:: :: |: |: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Gp 806 WSNCKWSNANCKYTTTAAATCNCKYTGCCYTYTGCCANCKYTTNACWNSGTATN 747

-62' tbtggcarytbaayggnmgdytbgartavtygytbacqä/mqdatgaattytgayathc 121
:
746 ACGARYTCNCCNGCYGTYTTCGCCACACNARRTANGGCRTTTATCTCTTCATY 687

Db 122 cngargarathaarcgncrcarttycraargargaycgngcnymbacnathtayg 181
 :
 596 tcvctgcancncba vvcnabnsdpcytnabccanabna cnsvtvtvttnarbtcncky 627

Db 182 aratggtbcaraayathtttgcnahtttymgdcargywvnnwnnwnnacngntggaaag 241
:: : : : ::|: : : : : : : : : :
c 526 tcnwsancccttnadavttttatanaacncttttcvtcttcacnaatnatattttctct 567

[illegible]

Db 300 ngtntbgararaarbtbgaraargayttacnmgdgngaarytbatgwnwnnyt 359

```
Db      360 bcayytbaarmgdtaytayqgmgdathybtbcaytavytbaargcnaargartaywnnca 419
```

Db 420 ycygcntggacnathgtngdgtngarathymbgd 456
:|::: | |::: : ::: :::

RESULT 10

ID	N30031	Standard; DNA; 301 BF.
AC	N50031;	
DT	04-SEP-1991	(first entry)
DE	Sequence encoding new modified human beta interferon polypeptides	

DE RNA 448.
KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.
OS Homo sapiens

FH	Key	Location/Qualifiers
FT	CDS	1..501
FT	/*tag= a	
DN	FD-15002	

PD 11-DEC-1985.
PF 17-MAY-1985; 105750.
PR 17-MAY-1984; GB-012564.

PI Bell LD, Boseley PG, Porter AG;
WPI: 85-311944/50.
DR P-PSDB; P50030.

PT plasmid transformed bacteria, with improved antiviral,

PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 21, page 40; 71pp; English.
CC Compared with interferon beta prep. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they
CC have higher therapeutic index; improved stability against microbial
CC breakdown during synthesis; and better in vivo solubility and
CC stability. They are also easier to recover from incubation mixts.
SQ Sequence 501 Bp; 110 A; 30 C; 69 G; 80 T;

Query Match	7.3%;	Score 50;	DB 3;	Length 501;
Best Local Similarity	15.3%;	Pred. No. 4.27e-05;		
Matches	69;	Conservative 13;	Mismatches 242;	Indels 4;
				Gaps 3;

```
Db 53 araarytbytbgcarytbaayggnmgdytbgartaytgyytbaargymgdaartayt 112
|| :: ||| : | : : : ||| : | : || : : | : | :
Cc 668 abmspsctvtvabccanabnacnasytmvttwarrtcncykvtcnwfnacnccrttnary 609
```

Db 113 tyggnttyccncargagarttgyggnnaaycarttycaragaargayngenyta 172
| : : : | : | : | : : :
Ca 508 tttgtatataattcttgtagtcvtgcacnactcgaanlaaaatvtnnncttgtcnckdpttta 549

[illegible]

Db 233 gntggaaygaracnathgtngaraaytbytbgcnaaygtntaycaycarathaaycayy 292

```
Db      293 tbaacngtnybtbgargara--arytbgaraa--rgargayttyacnmgdggnaarytba 349
```

Db 350 tgvnnwnnytbcayytbaarmgdtaytayvgngmgdathytbcaytayytbaargcnaarg 409
| : : : ||| : : : | : : : : ||| : : : | : : :

D_b 410 artaywnncaytgygentggacnathgtnmgdngatgathytbmgsdaaytttyattyya 469
 :: : :: | : : : | : ||:: | ! : : :

CP 306 TCTTGGGCGTCACAGCAATGCACGCGTTCCGTAAGGTTCATTGCAATCATCATTGAAAT 273

Ddb 470 thaaymgdytbacnggntay-ytbgdaayt 499
 :|: |: :|: :|: :|||

```

CP      Z4B  CANACNGGCGCAANARNW5RCANARROA11  Z10
RESULT  11

```

ID	N50023 standard; DNA; 501 BP.
AC	N50023;
DT	04-SEP-1991 (first entry)
DE	Sequence encoding non modified human beta interferon polypeptide

DE IFNX 416.
KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.

	Key	Location/Qualifiers
FH	CDS	1..501
FT	/tag= a	..
FT		
FT		

PD 11-DEC-1985.
 PF 17-MAY-1985; 105750.
 PR 17-MAY-1984; GB-012564.

PI	Bell LD, Boleley PG, Porter AG;
DR	WPI; 85-311944/50.
DR	P-PSDB: P50022.
PT	New modified human beta interferon polypeptide(s) - prep.d. by
PT	plasmid transformed bacteria, with improved antiviral,
PT	anti-proliferative and immune regulating actions
PS	Claim 28; Chart 2a, page 32; 71pp; English.
CC	Compared with interferon beta prep.d. by recombinant methods, the
CC	INFs of the invention are more active and have different affinities
CC	for cell surface receptors (allowing selective targeting); they
CC	have higher therapeutic index; improved stability against microbial
CC	breakdown during synthesis; and better in vivo solubility and
CC	stability. They are also easier to recover from incubation mixts.
SQ	Sequence 501 BP; 107 A; 31 C; 69 G; 80 T;
	Query Match 7.3%; Score 50; DB 3; Length 501;
	Best Local Similarity 14.9%; Pred. No. 4.27e-05;
	Matches 68; Conservative 138; Mismatches 246; Indels 4; Gaps 3;
Ddb	48 rwnccaraarybybtggcarycbmmvgngmgdytbgartaytyybaargaymgdca 107 :: : : : : : : : : : : : : : : : :
Cp	673 AYTGNAWRNSRTCYYTNARCCANARNACNWSYTTYTTNARRTCNKCYTCNWSNACNCRT 614
Ddb	108 ygaytytgnttyccncargarttygagynaacarttycarcargaygcngc 167 : : : : : : : : : : : : : : : : :
Cp	613 TNARYTTRTANAICYTTNWSYTYNGTYTCNARNATNATYTTNGTRTCCKRTTNART 554
Ddb	168 nytbacnathaygaratgytbcaraayathctygcnathctymgdcardgaywmnwnn 227 : : : : : : : : : : : : : : : : :
Cp	553 ANGNTAYCTYTNNACYTTNATYTTNARNCGAARTCRITYTTRCANRNGTYTCANNA 494
Ddb	228 nacngntggaaygaraecnathgtngaraaytybytblgcnaaqyntaycaycarathaa 287 : : : : : : : : : : : : : : : : :
Cp	493 TRTCRTTRTCRTCTRTTYTTRYTTTCANGCYTCRCANACYTTGNGGCYTCYTCNG 434
Ddb	288 ycaayytbaarcangtnybtargata--arytbgaraar-gargaytytcnmgdgnaa 344 : : : : : : : : : : : : : : : : :
Cp	433 TNGCNGGNARNARRTCRTCNWSNWSNGNARNGGNATRCANARRCTRTTCYTGNGGRA 374
Ddb	345 rytbatgwmmwnnytbcaayytbaarmgdktaytaggmngdathytbcaytayybaargc 404 : : : : : : : : : : : : : : : : :
Cp	373 ANCKRTCBAYTCNARCATRCTNGGCCGANGRAANCBAANGCWNSCATWACNGNGCRC 314
Ddb	405 naargattaymncaytygvngntggacnathgtmngdgtngrathctymbgdaaytyta 464 : : : : : : : : : : : : : : : : :
Cp	313 ANCKRTCYTNACYTCNACRCANRNWSTGRGANGGYTGNAITNGTYTCTCNARTCRT 254
Ddb	465 ytythaaymgdytbacngntay-ytbgmdaayt 499 : : : : : : : : : : : : : : : : :
Cp	253 CNARCNACNGNGGCRANARNMWSNCAVAAAYT 218

RESULT	12	
ID	N50029	standard; DNA; 501 BP.
AC	N50029;	
DT	04-SEP-1991	(first entry)
DE	Sequence encoding new modified human beta interferon polypeptides	
DE	IFN γ 446.	
KW	Antiviral;	cell growth regulator; immune system regulator;
KW	antiproliferative;	ss.
OS	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	CDS	1..501
FT	/*tag= a	

PN	EP-163993-A.
PD	11-DEC-1985.
PF	17-MAY-1985; 105750.
PR	17-MAY-1984; GB-012564.
PS	(SEAR) SEARLE G D & CO.
PI	Bell LD, Boseley PC, Porter AG;
PA	WPI; 85-311944/50.
DR	p-PSDB; P50028.
PT	New modified human beta interferon polypeptide(s) - prepd. by
DD	plasmid transformed bacteria, with improved antiviral,
PPT	anti-proliferative and immune regulating actions
PPT	Claim 28; Chart 2g, page 38; 71pp; English.
PS	CC Compared with Interferon beta prepd. by recombinant methods, the
CC	INFs of the invention are more active and have different affinities
CC	for cell surface receptors (allowing selective targeting); they
CC	have higher therapeutic index; improved stability against microbial
CC	breakdown during synthesis; and better in vivo solubility and
CC	stability. They are also easier to recover from incubation mixts.
CC	Sequence 501 BP; 112 A; 31 C; 69 G; 79 T;
SQ	
	Query Match: 7.3%; Score 50; DB 3; Length 501;
	Best Local Similarity 16.0%; Prec. No. 4.27e-05;
	Matches 72; Conservative 133; Mismatches 242; Indels 4; Gaps 3;
D b	53 araarybybtbgcarytbaayggumdytbgaratyggtbaargaymgdatgaayt 112 : : : : : : : : : : : : :
C p	668 ARNWRSTCYTNARCANARNACNWSYTYTNNARTCNCKYTCNWSNACNCRTNARY 609
D b	113 tygnathccngargarathaarcaargncarcarttycaraargagvgcngcnycba 172 : : : : : : : : : : : : : :
C p	608 TTRTANATNGYTNTWSYTTNGYTCNARNATNATYTTTGTCNCRTNATRANGTN 549
D b	173 cnathtyagaratgytbcaayathittgcgnathittymgdcatcgaymwnnnnacng 232 : : : : : : : : : : : : : :
C p	548 ATVCTYTNACTYNATTNARNGCRAARTCRITTYTCANARNGTYYCCATNATRTC 489
D b	233 gntggaygaracnathgtngaraaytybtbgcnaaygtnataycarathaaycayy 292 : : : : : : : : : : : : : :
C p	488 TTTRCRTCRTYTTTTRTTTCANCAGCYTCRCANACYTTGCGCYTCYTCNGTCGN 429
D b	293 tbaaracngntnycbgarga--arytbogaraa--rgargaytyacmgdgmaarycba 349 : : : : : : : : : : : : : : :
C p	428 GGARNARRTRTCNWSNWSNGCNARGNATRCANARRTCRTTTCYTCNGGRAACKR 369
D b	350 tgwnwnnytbcaayrbarmgdtaytayggumgdathytbcayttbayrtbaargnaarg 409 : : : : : : : : : : : : : : :
C p	368 TCRCAYTONCARATRTCCGCCANGRANCCRANGCNWSCATNACNGGCGRCANCR 309
D b	410 artatwmcaytygcntggacnathgtmngdgtngaratthybtmgdaaytyttaytya 469 : : : : : : : : : : : : : : :
C p	308 TCYTNACTYGNACRCANARNWGRTRGANGGYTGATNGTYTCRTONARTCRTCNAIR 249
D b	470 thaaymgdytbacngntay-ytbmgdaayt 499 : : : : : : : : : : : : : : : :
C p	248 CANACNGGCRANARNWSCRANARAAYT 218

RESULT 13
ID N50024 standard; DNA; 501 BP.
AC N50024;
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE DE
DE IFN 417.
KW Antiviral; cell growth regulator; immune system regulator;

KW	antiproliferative; ss.	
OS	Homo sapiens.	
Key	Location/Qualifiers	
FT	1..501	
FT	/*tag= a	
FT	EP-163993-A.	
PD	11-DEC-1985.	
PF	17-MAY-1985; 105750.	
PPR	17-MAY-1984; GB-012564.	
PA	(SEAR) SEARLE G D & CO.	
PI	Bell LD, Boseley PG, Porter AG;	
DR	WPI; 85-311944/50.	
DR	P-PSDB; p50023.	
PPT	New modified human beta interferon polypeptide(s) - prepd. by	
PPT	plasmid transformed bacteria, with improved antiviral,	
PPT	anti-proliferative and immune regulating actions	
PS	Claim 2b, page 33; 71pp; English.	
CC	Compared with interferon beta prepd. by recombinant methods, the	
CC	INFs of the invention are more active and have different affinities	
CC	for cell surface receptors (allowing selective targeting); they	
CC	have higher therapeutic index; improved stability against microbial	
CC	breakdown during synthesis; and better in vivo solubility and	
CC	stability. They are also easier to recover from incubation mixts.	
SQ	Sequence 501 BP; 110 A; 32 C; 66 G; 81 T;	
Query Match	7.1%; Score 49; DB 3; Length 501;	
Best Local Similarity	16.0%; Pred. No. 8.62e-05;	
Matches	50; Conservative 94; Mismatches 168; Indels 1; Gaps 1;	
Db	166 gcnymbacnathataygaratgytbcaraayathttgcnathttymgdcargaywnwnn 225	
Qy	410 SNMSGAYCAYTYNTYTCNGCNCACGARGAGCCNCAARGTNTGYGARGCNTGYAARA 469	
Db	226 wnnacngntggaaygaracnathgtngaraaytbytbcgnaaygtnt-taycaycarat 284	
Qy	470 AYAAARAAYGAYGAYAAAYGAYATNATGACACNTNTGYAARAAYZAYTTYGCNTYNA 529	
Db	285 haaycayytbaaracngntnycbgargaraarytbgaraargaytytcaacmgdgnaa 344	
Qy	530 ARATNARGTNAARGARATNACNTAYATNAAVMGNGAYACVAARATNATNTYNGARACNA 589	
Db	345 rytcbatgwnwnnnytbcaytbaarmgdattaytggngmdathytbcaytaytbaatgc 404	
Qy	590 ARHWAARACNATNTAYATNTYNAAYTCGNTWNSGARMGNGAYTYTNAARAARWSNGINY 649	
Db	405 naargattaywnncattgygcntgacnathgtmgdgtngarathytbmgdaaettyta 464	
Qy	650 TNGTGYTNAARGAYWSNYTNTNARTGYACNTGYCARGARATGAAYGAYATNAYGCNCNT 709	
Db	465 ytttyathaaumgd 477	
Qy	710 AYTGTGNTATGGC 722	

RESULT	14	
ID	N50032	standard; DNA; 501 BP.
AC	N50032;	
DT	04-SEP-1991	(first entry)
DE	Sequence encoding new modified human beta interferon polypeptides	
DE	INX 449.	
DE	Antiviral; cell growth regulator; immune system regulator;	
KW	antiproliferative; ss.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers

FT	CDS		1..501	
FT		/t=tag= a		
PN	EP-163993-A.			
PD	11-DEC-1985.			
PE	17-MAY-1985; 105750.			
PF	17-MAY-1984; GB-012564.			
PR	(SEAR) SEARLE G D & CO.			
PA	Bell LD, Boseley PC, Porter AG;			
PI	WPI; 85-311944/50.			
PJ	P-PSDB; P50031.			
DR	New modified human beta interferon polypeptide(s) - prepd. by			
PT	plasmid transformed bacteria, with improved antiviral,			
PT	anti-proliferative and immune regulating actions			
PS	Claim 28; Chart 23, page 41; 71pp; English.			
CC	Compared with interferon beta prep. by recombinant methods, the			
CC	INFs of the invention are more active and have different affinities			
CC	for cell surface receptors (allowing selective targeting); they			
CC	have higher therapeutic index; improved stability against microbial			
CC	breakdown during synthesis; and better in vivo solubility and			
CC	stability. They are also easier to recover from incubation mixts.			
SQ	Sequence 501 BP; 108 A; 30 C; 72 G; 79 T;			
	Query Match	71.1%; Score 49; DB 3; Length 501;		
	Best Local Similarity	15.3%; Pred. No. 8.62e-05;		
	Matches	69; Conservative 135; Mismatches 243; Indels 4; Gaps 3;		
Db	53	araarytybtgdcarytbaayggmgdytbgartaygyvtbaargaymcdgayv	112	
		: : : : : : : : : : : : : : : : : : :		
Cp	668	ARNWSRCTYTNARCACNARNACNWSTYYTNTNARRTCNKCYTCNWSNACNCRTNARY	609	
Db	113	tygnntcyccngngargarttyagygnaaycarttcyaaragaygcngcnybta	172	
		: : : : : : : : : : : : : : : : : : :		
Cp	608	TTRTANATNGTYTNWSYTNGTYTCNARNATTYTTNCTRCKRRTNATRYANGTN	549	
Db	173	cnathtaygarctygtbcaraayathitgcnathnattymgdcargaywnnnwmaacng	232	
		: : : : : : : : : : : : : : : : : : :		
Cp	548	ATYCTYNACTYNATTYTNWARGCBAARTCRITYTTCANARNGITYTCCATNATRCR	489	
Db	233	qntggaagaracnathgtingaraaytybtbgcnaaygtnatcaycarathaacyav	292	
		: : : : : : : : : : : : : : : : : : : : :		
Cp	488	TTRCTCRTCTRTYTTTRTYTTRTCANGCYTCRANACYTTGGGCGYTCYTCGTNGCN	429	
Db	293	tbaarcngtnybtbgarga--arytbgaraa--rgargaytiyacmgdgnaaarytba	349	
		: : : : : : : : : : : : : : : : : : : :		
Cp	428	GGNARRRTGRTCWNWSNGNRRNGGATRCANARRTCRTTTRTCYTGNGGRAACKR	369	
Db	350	tgwnnnynytbcayvtbaarmgdcatytaggmugdathytbcaytayvbaargcnaarg	409	
		: : : : : : : : : : : : : : : : : : :		
Cp	368	TCRCATYCNACATRCITCGGCCGNGGRANCCBAAACGNWSCATNACNGGNGCRCANCKR	309	
Db	410	artaywnncaytygcntgagacnathqtgmmdgdtngaratthytblmgdaaytytatty	469	
		: : : : : : : : : : : : : : : : : : : : :		
Cp	308	TCYTTNACYTCNACRCANARNWSRTGRCAAGGYTGNAINTGYTTCRTONARTCTCNARR	249	
Db	470	thaaymgdytbacngntay-ytblmgdaayt	499	
		: : : : : : : : : : : : : : : : : : : :		
Cp	248	CANACNGGNGCRCANARNWSRCANARRAAYT	218	

RESULT 15
ID N50025 standard; DNA; 501 BP.
AC N50025:
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides

